# ES索引库操作

## 前言

我们先回顾一下，我们知道在es中有索引库的概念index，索引库就相当于数据库中的表table，而是索引库里有很多的文档document，文档就相当于数据库表中一行一行的记录row，我们知道，数据库要先创建表，才能去添加记录，es也一样，先得有了索引库，才能往里面去添加文档，因此，我们要先来学习一下索引库的操作，再去学习文档的操作。

要想去创建一个索引库，就像去建表一样：在建表语句里要去指定Schema，即是对字段Cloumn的约束。同样，在创建索引库时，需要去指定对应的mapping 映射：是对文档的约束，其实更好理解成是对字段Feild的约束。

因此先来看看 这些mapping约束和数据库约束Schema有什么区别？Mapping到底要定义什么东西，之后在看看怎么实现索引库的CRUD。

## Mapping隐射常见属性

Mapping映射是对索引库中文档字段的约束（比如字段什么类型啊？是否要建立索引啊？是否要分词啊？如果分词的话使用什么分词器啊），我们知道ES中的文档是JSON风格的，如下图：



### Type：字段的数据类型关键字，常见的简单类型有：

第一个约束的就是字段的数据类型，因为你写一个文档，文档的字段一定都有对用的数据类型啊！文档字段数据类型的约束主要是针对JSON里面的数据类型的，JSON里面要么是数字、要么是布尔值、要么是字符串、要么是数组、要么是对象嵌套，这些类型对应到ES中都有对应的类型去进行约束！

#### 1：字符串：

ES中把字符串分成两种，一种是text，一种是keyword，第一种text是可分词的文本，比如上图的info字段（个人信息介绍），是一长串、一句话，或者我们聊过的商品标题信息，都是一长串的文本，都是由多个词语组成的，像往往这种内容，往往可以分开的，即text，是可分词的，可拆分的！

第二类是keyword，是精确值，比如品牌、国家、IP地址、邮箱，作为一个整体才有意义，不管多长，拆分就没有意义了，比如：美特斯邦威、北爱尔兰及大不列颠联合王国等，都不能拆，是完整不可分割的整体。所以keyword的关键就是字段不能拆，只有合在一起才有意义！

text（可分词的文本，例如：个人介绍、商品标题等）、keyword（精确值，例如：品牌、国家、IP地址、邮箱等）。所以，以后去设置字段映射时，一定要想清楚这个字段需不需要拆分，不需要就是keyword，需要就是text。

#### 2：数值类型：long、integer、short、byte、double、float

是不是和JAVA特别像，因为ES底层本来就是JAVA实现的！

#### 3：布尔：boolean

#### 4：日期：date

#### 5：对象：object

#### 注意

es中是没有数组这中类型的，但是，允许你某个类型的字段有多个值，比如说，字段是Boolean类型，但是这个字段有多个值，没问题，可以的。也就是说，只需要关注数组里元素的类型，如上图中的字段score的类型就是float或double！

### Index：是否创建索引

是否创建索引就是字面意思，即要不要创建倒排索引，默认为true，则会为字段创建倒排索引，那么就可以参与搜索了，实际开发中，并不需要每个字段都需要去搜索，比如商品图片url地址、邮箱，没有搜索意义，那么因此，在创建字段映射时，一定要判断这个字段是否参与不参与搜索，不参与要设置成false。（注意：是否索引和是否分词没有必然的关系，但是要分词了的字段一般都是要做索引，但做索引的字段不一定分词）

#### 注意点1：index的意思就是要不要为这个字段建立倒排索引

字段的Index默认值是true，也就是说你在做字段映射时，如果你不去修改index，那默认所有的字段都会被创建倒排索引！还有记住了，是否索引和是否分此没有必然关系，没分词的字段，赵柯可以创建倒排索引，但是对一个字段分词，index可以肯定为true，即对文档字段分了词，就是要对字段建立倒排索引！

#### 注意点2：index值比较难判断的是对于字符串类型的字段

#### 注意点3：index取true还是false得关键是看这个字段要不要参与搜索

#### 注意点4：index取值跟字符串类型的字段是text还是keyword无关，意思就是说字段为keyword，可能index为false，也可能是true，关键还是要看这个字段参不参与搜素！

### Analyzer：使用哪种分词器，

用来对内容做分词的，可想而知，这个属性用得比较少，因为在上面的数据类型里，只有一个数据类型需要分词：text，其他类型都无需分词，因此，这个**Analyzer属性是结合text类型一起使用**，值就是分词器的名称。

还是这句话：当一个字段的type是text，同时分了词analyzer，那么这个字段的index必然是true，不然你分词干嘛！而且要不要分词一般只是争对与type为text的字敦，其他字段都没有分词的必要啊！

### Properties：该字段的子字段，即object对象类型中子属性



### 小总结：

#### Mapping的常见属性有哪些：

Type：数据类型；

Index：是否索引；

Analyzer：分词器

Properties：子字段

#### Type常见的字段有哪些：

字符串：text、keyword

数字：long、integer、short、byte、double、float

布尔：Boolean

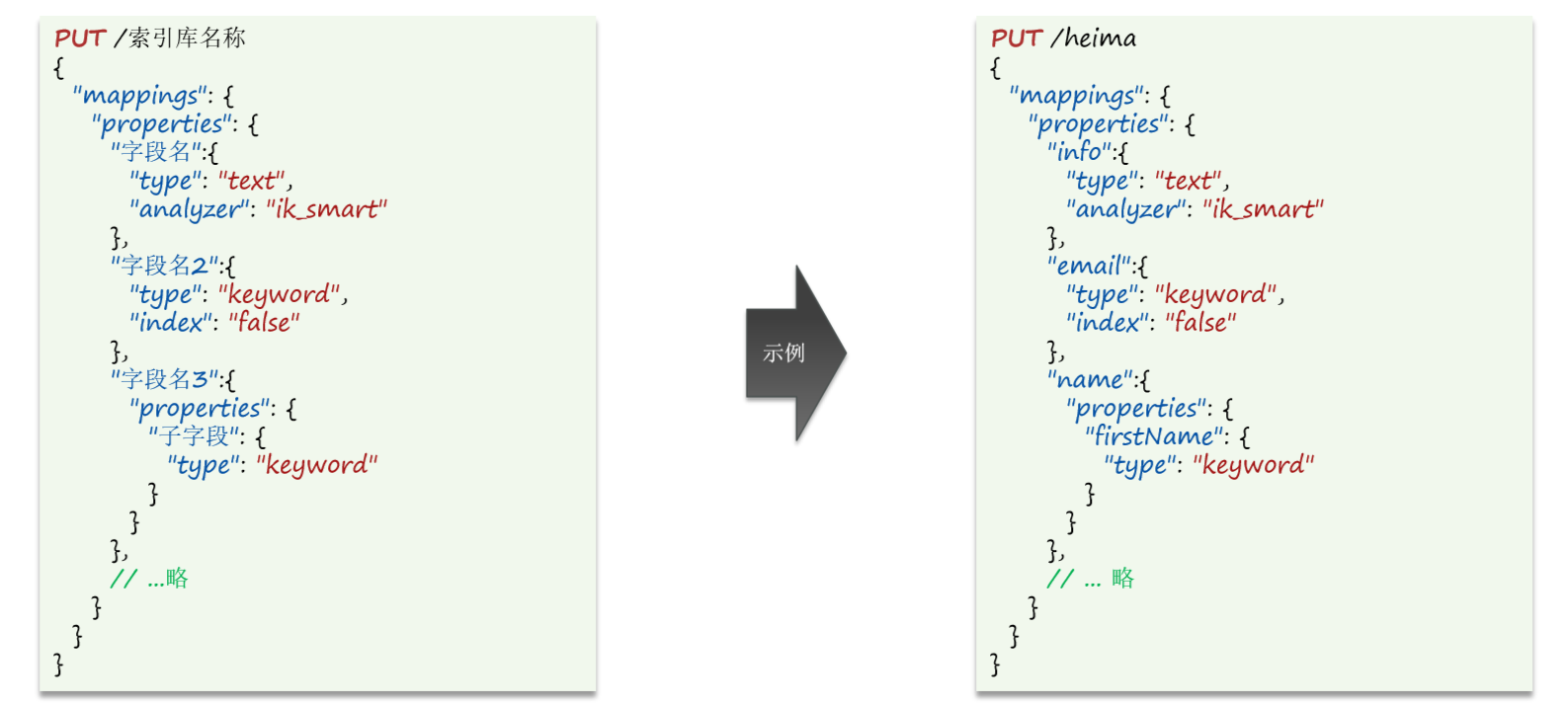
日期：date

对象：object

## 创建索引库

上面已经了解了mapping里边常见的属性，下边我们创建索引库的时候就需要用到这些属性，但是创建索引库的DSL语法怎么写呢？请看以下说明：

ES通过Restful请求操作索引库、文档，请求内容用DSL语句来表示。传概念索引库和mapping的DSL语法如下：



### 通用格式：

**PUT /索引库名称**

**{**

**“mappings”: {**

**“properties”: {**

**“字段名”: {**

**“type”: “XXX”,**

**“analyzer”: “XXX”,**

**“index”: “ture or false”**

**},**

**“字段名”: {**

**“properties”: {**

**“子字段名”: {**

**“type”: “XXX”,**

**“analyzer”: “XXX”,**

**“index”: “ture or false”**

**}**

**}**

**}**

**}**

**}**

**}**

### 创建heima索引库

**PUT /heima**

**{**

**"mappings": {**

**"properties": {**

**"info": {**

**"type": "text",**

**"analyzer": "ik\_smart"**

**},**

**"email": {**

**"type": "keyword",**

**"index": false**

**},**

**"name": {**

**"type": "object",**

**"properties": {**

**"firstName": {**

**"type": "keyword"**

**},**

**"lastName": {**

**"type": "keyword"**

**}**

**}**

**}**

**}**

**}**

**}**

## 查询索引库：GET /索引库名

## 删除索引库：DELETE /索引库名

## 修改索引库

事实上，在es中，索引库是不允许修改的（注意，修改的是索引库，不是修改文档，就是修改文档是可以的），因为索引库创建完了以后，她的数据结构，即索引库的mapping映射都已经定义好了，我们的ES会基于mapping映射去创建倒排索引，那么如果说，你要去修改索引库的一个字段的话，就会导致原有的整个倒排索引就会彻底失效，这样一来带来的影响是特别巨大的，于是在ES里面，是禁止修改索引库的，这点和数据库不一样，数据库中，一张表创建完了，是可以去修改表结构的（理论上可以，但是在生产环境下，数据库也是避免去修改的表结构的）。

但是，ES虽然禁止你直接去修改索引库原有的字段，但是允许你去增加新字段，增加字段这也是属于修改嘛

千万注意，新字段一定是一个全新的字段名，一定不能和以前的字段重复，否则的话，你就是在改一个存在的字段，会报错！

### 通用格式

**PUT /索引库名/\_mapping**

**{**

**“properties”: {**

**“新字段名”: {**

**“type”: “XXX”**

**}**

**}  
}**

### 给索引库heima新增age字段

**PUT /heima/\_mapping**

**{**

**"properties": {**

**"age": {**

**"type": "integer"**

**}**

**}**

**}**

## 总结

索引库操作有哪些？

1：创建索引库：POST /索引库名

2：查询索引库：GET /索引库名

3：删除索引库：DELETE /索引库名

4：添加字段：PUT /索引库名/\_mapping