

## 全文搜索 API 及维护操作

### 一、maven 引用 jar 包

将 jar 包放入 lib 目录，在 pom 添加

```
<dependency>
  <groupId>com.laoshan</groupId> <!--自定义-->
  <artifactId>es</artifactId> <!--自定义-->
  <version>1.0</version> <!--自定义-->
  <scope>system</scope> <!--system，类似 provided，需要显式提供依赖的
jar 以后，Maven 就不会在 Repository 中查找它-->
  <systemPath>${basedir}/lib/elasticsearch-1.0.jar</systemPath>
<!--项目根目录下的 lib 文件夹下-->
</dependency>
```

同时添加如下依赖

```
<dependency>
  <groupId>com.alibaba</groupId>
  <artifactId>easyexcel</artifactId>
  <version>2.2.6</version>
</dependency>
<dependency>
  <groupId>org.elasticsearch</groupId>
  <artifactId>elasticsearch</artifactId>
  <version>7.8.0</version>
</dependency>
<dependency>
  <groupId>org.elasticsearch.client</groupId>
  <artifactId>elasticsearch-rest-high-level-client</artifactId>
  <version>7.8.0</version>
</dependency>
<dependency>
  <groupId>com.alibaba</groupId>
  <artifactId>fastjson</artifactId>
  <version>1.2.79</version>
</dependency>
<dependency>
  <groupId>org.projectlombok</groupId>
  <artifactId>lombok</artifactId>
  <version>1.18.22</version>
  <scope>provided</scope>
</dependency>
```

## 二、方法接口

### 1、创建客户端对象

```
ES_Client es_Client=new ES_Client("localhost", 9200, "http");
```

### 2、excel 导入 es 方法

```
es_Client.insertDocToIndex("path", "laoshan_policy");
```

参数一：文件目录

参数二：索引名

说明：本方法为增量添加，不修改文档 id 的前提下，可自动更新 id 已存在的文档内容，若索引中不存在当前 id，默认为添加新文档。

### 3、字段查询方法（分词匹配，非全词匹配）

```
①es_Client.multiRequest("企业","policy_laoshan");
```

```
②es_Client.titleRequest("企业","policy_laoshan");
```

```
③es_Client.allRequest("policy_laoshan");
```

①多字段查询，包含 title, content, outline3 个字段

②只查询标题，只查询标题

③无条件列出所有文档

返回值：

```
System.out.println(es_Client.getResultForJson());
```

contentHighLight: content 字段的匹配值

titleHighLight: title 字段的匹配值

outlineHighLight: outline 字段的匹配值

url: 备用字段，可放 url

\*\*\*所有字段匹配值均已加入高亮红色字体标签\*\*\*

["contentHighLight": "<p>为加快推进<span style='color:red'>企业</span>家队伍建设，最大限度激发和

释 放 <span style='color:red'>企 业 </span> 家 创 新 创 业 活 力 ， 鼓 励 <span style='color:red'>企业</span>家

带领<span style='color:red'>企业</span>做大做强，加快推动新旧动能转换，根据《崂山区关于鼓励吸引高

层次人才创新创业加快推动新旧动能转换的实施办法》（崂发〔,"titleHighLight": "青岛市崂山区 “<span s

tyl='color:red'>企业</span>家培养计划” 实施细则", "outlineHighLight": "<p>（一）新增规模以上<span

style='color:red'>企业</span><span style='color:red'>企业</span>家人才资金奖励政策</p><p><b>内容

概要：</b>对纳入国家统计局联网直报系统的新增规模以上<span style='color:red'>企业</span></span>（不含拆分

）， 给 予 <span style='color:red'>企 业 </span> 家 最 高 100 万 元 奖 励 。

```
","url":"http://about:blank"]}]
```

#### 4、获取本次查询相关信息

```
①System.out.println(es_Client.resultInfoPojo.getQueryCount());
②System.out.println(es_Client.resultInfoPojo.getQueryTime());
③System.out.println(es_Client.resultInfoPojo.getQueryString());
```

①获取本次查询符合条件的文档数量，long 类型

②获取本次查询的耗时，long 类型，单位 ms

③获取本次查询的所有结果，以 json 格式导出，本结果无高亮标签。

```
{"begin_date":"2019-07-30","code":"青崂工信发〔2019〕28号","content":"<p>为加快推进企业家队伍建设，最大限度激发和释放企业家创新创业活力，鼓励企业家带领企业做大做强，加快推动新旧动能转换，根据《崂山区关于鼓励吸引高层次人才创新创业加快推动新旧动能转换的实施办法》(崂发〔2017〕22号)相关规定(省略).....",
"create_date":"2021-07-02","end_date":"2022-08-31","outline":"<p>(一)新增规模以上企业企业家人才资金奖励政策</p><p><b>内容概要：</b>对纳入国家统计局联网直报系统的新增规模以上企业（境外“买壳”上),(省略.....)
"subject":"人才支持","title":"青岛市崂山区“企业家培养计划”实施细则","update_date":"2022-02-17"}
```

#### 5、聚合查询

API 设计还需当面商讨!

#### 6、断开连接

```
es_Client.DisconnectionES();
```

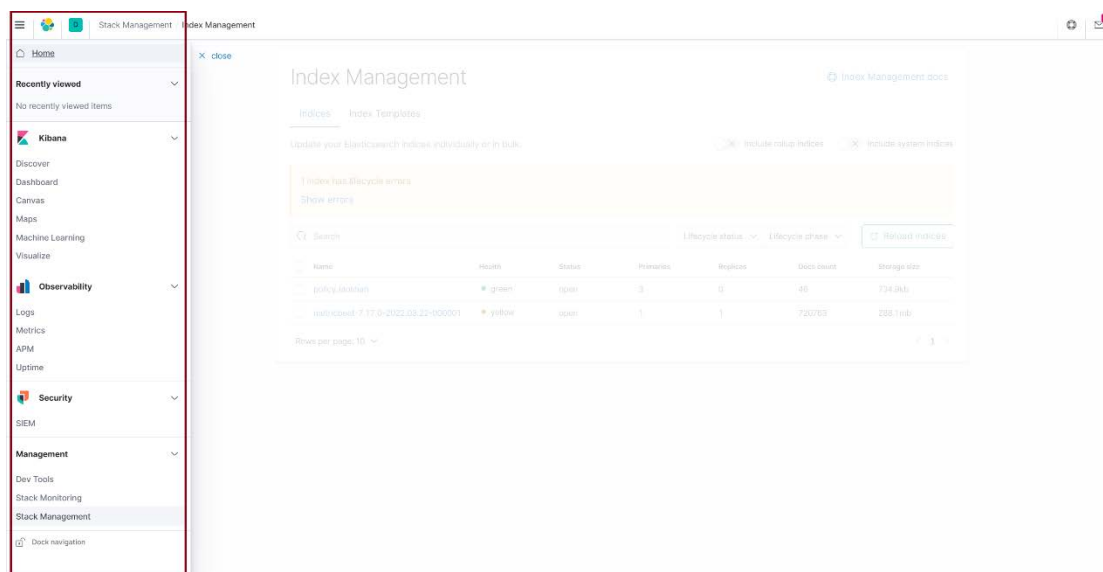
### 三、数据维护（kibana 应用）

服务器端启动 kibana (cd 进入 kibana 下 bin 文件夹)

./kibana

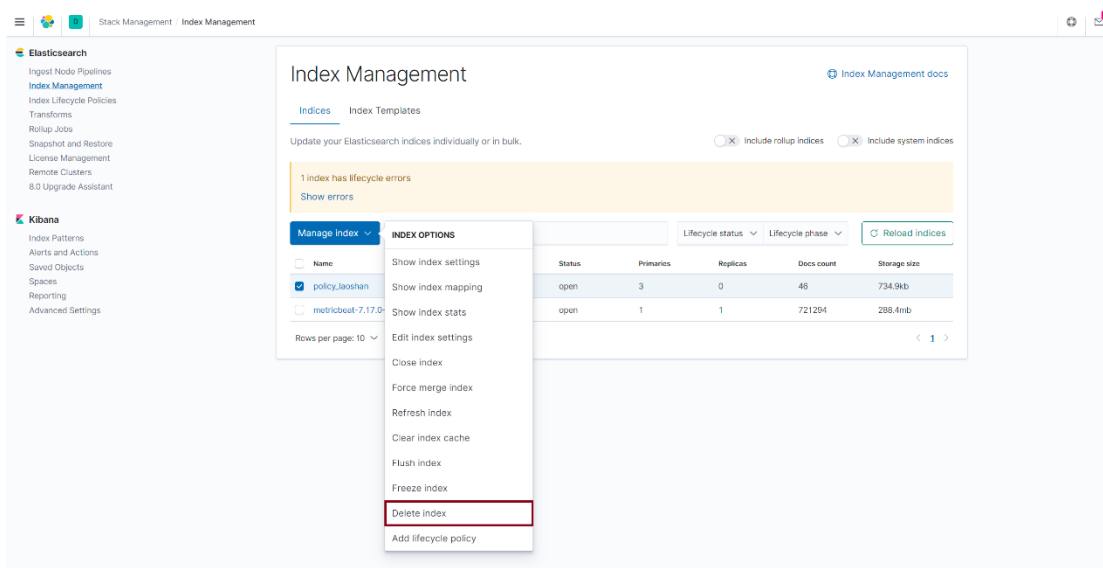
打开浏览器 <http://localhost:5601/>

以下操作都在左侧菜单完成



## 1、删除索引

选中菜单 Stack Management---Index Management---选中当前索引---Delete Index

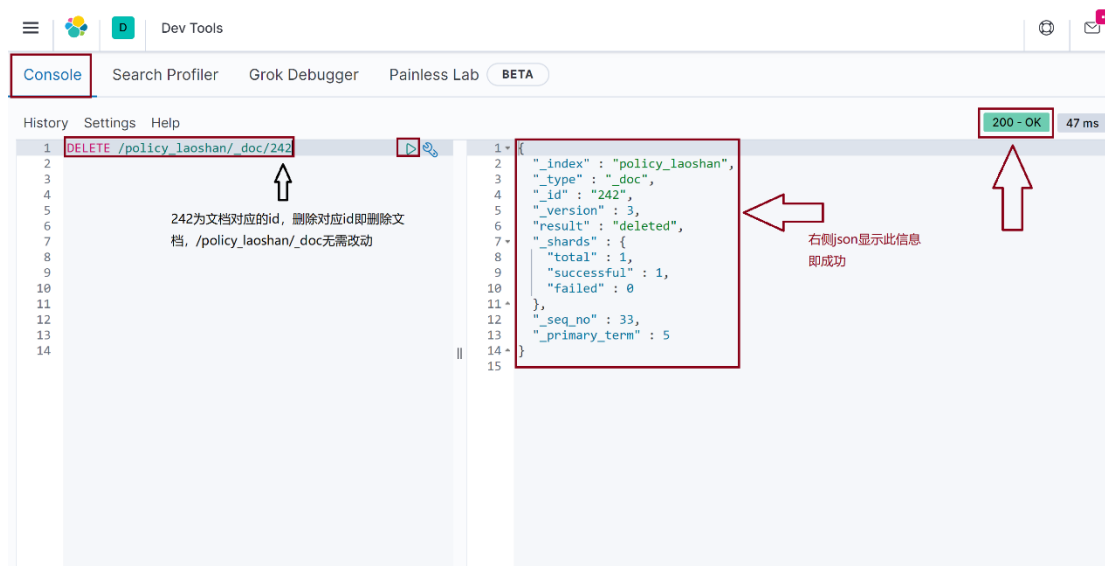


\*\*\*此方法会删除当前索引以及索引下所有数据，数据搜索会找不到索引，若恢复数据必须重新建立索引，非必要请勿直接删除索引，想删除部分数据请使用下列方法\*\*\*

## 2、删除数据

选中菜单 Dev Tools---Console---写入下列命令---点击右侧 运行

DELETE /policy\_laoshan/\_doc/{文档 id}



## 四、ES 配置

依据硬件环境，采用单主机 3 分片无副本架构。优点：查询速度快，查询时无需查找其他分片，节省查询时间。缺点：若主机宕机，数据即无法恢复，需按时本地备份数据。

因项目进度要求，暂采用 ES7.8 版本，后期可依据需求升级至最新版本