http://blog.csdn.net/lihang\_1994/article/details/72599449

为什么要用solr服务，为什么要用luncence？

问题提出：当我们访问购物网站的时候，我们可以根据我们随意所想的内容输入关键字就可以查询出相关的内容，这是怎么做到呢？这些随意的数据不可能是根据数据库的字段查询的，那是怎么查询出来的呢，为什么千奇百怪的关键字都可以查询出来呢？

答案就是全文检索工具的实现，luncence采用了词元匹配和切分词。举个例子：北京天安门------luncence切分词：北京 京天 天安 安门 等等这些分词。所以我们搜索的时候都可以检索到。

有一种分词器就是IKAnalyzer中文分词器，它有细粒度切分和智能切分，即根据某种智能算法。

这就使用solr的最大的好处：检索功能的实现。

使用步骤；

（１）solr服务器搭建，因为solr是用java开发的，所以需要jdk和tomcat。搭建部署

（２）搭建完成后，我们需要将要展示的字段引入solr的库中。配置spring与solr结合，工程启动的时候启动solr

（３）将数据库中的查询内容导入到solr索引库，这里使用的是solrj的客户端实现的。具体使用可以参考api

（４）建立搜索服务，供客户端调用。调用solr，查询内容，这中间有分页功能的实现。solr高亮显示的实现。

（５）客户端接收页面的请求参数，调用搜索服务，进行搜索。

业务字段判断标准：

1、在搜索时是否需要在此字段上进行搜索。例如：商品名称、商品的卖点、商品的描述

（这些相当于将标签给了solr，导入商品数据后，solr对这些字段的对应的商品的具体内容进行分词切分，然后，我们就可以搜索到相关内容了）

2、后续的业务是否需要用到此字段。例如：商品id。

需要用到的字段：

1、商品id

2、商品title

3、卖点

4、价格

5、商品图片

6、商品分类名称

7、商品描述

Solr中的业务字段：

1、id——》商品id

其他的对应字段创建solr的字段。

<field name="item\_title" type="text\_ik" indexed="true" stored="true"/>

<field name="item\_sell\_point" type="text\_ik" indexed="true" stored="true"/>

<field name="item\_price" type="long" indexed="true" stored="true"/>

<field name="item\_image" type="string" indexed="false" stored="true" />

<field name="item\_category\_name" type="string" indexed="true" stored="true" />

<field name="item\_desc" type="text\_ik" indexed="true" stored="false" />

<field name="item\_keywords" type="text\_ik" indexed="true" stored="false" multiValued="true"/>

<copyField source="item\_title" dest="item\_keywords"/>

<copyField source="item\_sell\_point" dest="item\_keywords"/>

<copyField source="item\_category\_name" dest="item\_keywords"/>

<copyField source="item\_desc" dest="item\_keywords"/>

Solr 是Apache下的一个顶级开源项目，采用Java开发，它是基于Lucene的全文搜索服务器。Solr提供了比Lucene更为丰富的查询语言，同时实现了可配置、可扩展，并对索引、搜索性能进行了优化。

Solr是一个全文检索服务器，只需要进行配置就可以实现全文检索服务。有效降低频繁访问数据库对数据库造成的压力。

第一步：将solr部署在linux系统下。

第二步：solrJ是solr的客户端，使用它需要依赖solrJ的jar包。

第三步：将数据库的内容添加到solr的索引库，这样查询就在索引库查询，而不是数据库了。

Controller层：