

## 使用 Apache Solr 检索数据

在《[Apache Solr 介绍及安装部署](#)》文章里面我简单地介绍了如何在 Linux 平台搭建单机版的 Solr 服务，而且我们已经创建了一个名为 iteblog 的 core，已经导入了相关的索引数据，接下来让我们来使用 Solr 检索这些数据。



如果想及时了解 Spark、Hadoop 或者 Hbase 相关的文章，欢迎关注微信公共帐号：iteblog\_hadoop

### 查询所有的数据

可以使用 `*:*` 通配符来查找所有的索引数据，默认返回前 10 条数据：

```
[root@iteblog.com /opt/solr-7.4.0]$ curl http://iteblog.com:8983/solr/iteblog/select?q=*:*
{
  "responseHeader":{
    "status":0,
    "QTime":0,
    "params":{
      "q":"*:*"}},
  "response":{"numFound":52,"start":0,"docs":[
    {
      "id":"0553573403",
      "cat":["book"],
      "name":["A Game of Thrones"],
      "price":[7.99],
      "inStock":[true],
      "author":["George R.R. Martin"],
      "series_t":"A Song of Ice and Fire",
```

```
"sequence_i":1,
"genre_s":"fantasy",
"name_str":["A Game of Thrones"],
"cat_str":["book"],
"author_str":["George R.R. Martin"],
"_version_":1606764103107346432},
.....
}
```

我们可以通过 rows 参数来设置一次返回的条数：

```
[root@iteblog.com /opt/solr-7.4.0]$ curl http://iteblog.com:8983/solr/iteblog/select?q=*:*W&rows=1
{
  "responseHeader":{
    "status":0,
    "QTime":1,
    "params":{
      "q":"*:*",
      "rows":"1"}},
  "response":{"numFound":52,"start":0,"docs":[
    {
      "id":"0553573403",
      "cat":["book"],
      "name":["A Game of Thrones"],
      "price":[7.99],
      "inStock":[true],
      "author":["George R.R. Martin"],
      "series_t":"A Song of Ice and Fire",
      "sequence_i":1,
      "genre_s":"fantasy",
      "name_str":["A Game of Thrones"],
      "cat_str":["book"],
      "author_str":["George R.R. Martin"],
      "_version_":1606764103107346432}}
  ]}
}
```

## 检索指定关键字的数据

查询所有的索引一般没啥意义，所以让我们查询某个关键字的数据吧：

```
[root@iteblog.com /opt/solr-7.4.0]$ curl http://iteblog.com:8983/solr/iteblog/select?q=electronics
{
  "responseHeader":{
    "zkConnected":true,
    "status":0,
    "QTime":6,
    "params":{"q":"electronics"}},
  "response":{"numFound":14,"start":0,"maxScore":1.5579545,"docs":[
    {
      "id":"IW-02",
      "name":"iPod & iPod Mini USB 2.0 Cable",
      "manu":"Belkin",
      "manu_id_s":"belkin",
      "cat":["electronics",
        "connector"],
      "features":["car power adapter for iPod, white"],
      "weight":2.0,
      "price":11.5,
      "price_c":"11.50,USD",
      "popularity":1,
      "inStock":false,
      "store":"37.7752,-122.4232",
      "manufacturedate_dt":"2006-02-14T23:55:59Z",
      "_version_":1574100232554151936,
      "price_c___l_ns":1150}}
  ],
  .....
}
```

上面查询了 electronics 关键字的数据。但是可能会有很多字段里面包含了这个关键字，如果我们想查找特定字段里面出现这个关键字的，那么我们可以这样写：

```
[root@iteblog.com /opt/solr]$ curl http://iteblog.com:8983/solr/iteblog/select?q=manu:electronics
{
  "responseHeader":{
    "status":0,
    "QTime":4,
    "params":{
```

```
"q":"manu:electronics"}},
"response":{"numFound":1,"start":0,"docs":[
{
  "id":"SP2514N",
  "name":"Samsung SpinPoint P120 SP2514N - hard drive - 250 GB - ATA-133",
  "manu":"Samsung Electronics Co. Ltd.",
  "manu_id_s":"samsung",
  "cat":["electronics",
    "hard drive"],
  "features":["7200RPM, 8MB cache, IDE Ultra ATA-133",
    "NoiseGuard, SilentSeek technology, Fluid Dynamic Bearing (FDB) motor"],
  "price":92.0,
  "price_c":"92.0,USD",
  "popularity":6,
  "inStock":true,
  "manufacturedate_dt":"2006-02-13T15:26:37Z",
  "store":"35.0752,-97.032",
  "_version_":1606843234767601664,
  "price_c___l_ns":9200}]
}}
```

上面例子我们只查找在 manu 里面出现 electronics 关键字的数据，如果想在其他字段里面查找关键字也可以这样写。

## 短语搜索

有时候我们需要搜索一个短语，在 Solr 可以用下面语句实现：

```
[root@iteblog.com /opt/solr]$ curl http://iteblog.com:8983/solr/iteblog/select?q=W"TFT+LCD
W"
{
  "responseHeader":{"
    "status":0,
    "QTime":0,
    "params":{"
      "q":"W"TFT LCDW""}},
  "response":{"numFound":2,"start":0,"docs":[
    {
      "id":"9885A004",
      "name":"Canon PowerShot SD500",
      "manu":"Canon Inc.",
```

```
"manu_id_s":"canon",
"cat":["electronics",
"camera"],
"features":["3x zoop, 7.1 megapixel Digital ELPH",
"movie clips up to 640x480 @30 fps",
"2.0W" TFT LCD, 118,000 pixels",
"built in flash, red-eye reduction"],
"includes":"32MB SD card, USB cable, AV cable, battery",
"weight":6.4,
"price":329.95,
"price_c":"329.95,USD",
"popularity":7,
"inStock":true,
"manufacturedate_dt":"2006-02-13T15:26:37Z",
"store":"45.19614,-93.90341",
"_version_":1606843238758481920,
"price_c___l_ns":32995},
{
"id":"MA147LL/A",
"name":"Apple 60 GB iPod with Video Playback Black",
"manu":"Apple Computer Inc.",
"manu_id_s":"apple",
"cat":["electronics",
"music"],
"features":["iTunes, Podcasts, Audiobooks",
"Stores up to 15,000 songs, 25,000 photos, or 150 hours of video",
"2.5-inch, 320x240 color TFT LCD display with LED backlight",
"Up to 20 hours of battery life",
"Plays AAC, MP3, WAV, AIFF, Audible, Apple Lossless, H.264 video",
"Notes, Calendar, Phone book, Hold button, Date display, Photo wallet, Built-in games, JP
EG photo playback, Upgradeable firmware, USB 2.0 compatibility, Playback speed control, Rec
hargeable capability, Battery level indication"],
"includes":"earbud headphones, USB cable",
"weight":5.5,
"price":399.0,
"price_c":"399.00,USD",
"popularity":10,
"inStock":true,
"store":"37.7752,-100.0232",
"manufacturedate_dt":"2005-10-12T08:00:00Z",
"_version_":1606843234808496128,
"price_c___l_ns":39900}
}}
```

上面的例子我们搜索了包含 TFT LCD 两个关键字的文档，请注意，两个单词必须使用双引号括起来，否则会搜索出只包含其中一个关键字的文档。

## 联合搜索 ( Combining Searches )

我们可能还会从文档中搜索既包含 A 又包含 B 的文档，在 Solr 里面可以这样写：

```
[root@iteblog.com /opt/solr]$ curl http://iteblog.com:8983/solr/iteblog/select?q=%2Bprinter%20%2Bcopier
{
  "responseHeader":{
    "status":0,
    "QTime":0,
    "params":{
      "q":"+printer +copier"}},
  "response":{"numFound":1,"start":0,"docs":[
    {
      "id":"0579B002",
      "name":"Canon PIXMA MP500 All-In-One Photo Printer",
      "manu":"Canon Inc.",
      "manu_id_s":"canon",
      "cat":["electronics",
        "multifunction printer",
        "printer",
        "scanner",
        "copier"],
      "features":["Multifunction ink-jet color photo printer",
        "Flatbed scanner, optical scan resolution of 1,200 x 2,400 dpi",
        "2.5W" color LCD preview screen",
        "Duplex Copying",
        "Printing speed up to 29ppm black, 19ppm color",
        "Hi-Speed USB",
        "memory card: CompactFlash, Micro Drive, SmartMedia, Memory Stick, Memory Stick Pr
o, SD Card, and MultiMediaCard"],
      "weight":352.0,
      "price":179.99,
      "price_c":"179.99,USD",
      "popularity":6,
      "inStock":true,
      "store":"45.19214,-93.89941",
      "_version_":1606843234924888064,
      "price_c___l_ns":17999}]
  }}
}
```

上面的含义是从文档里面查找既包含 printer 又包含 copier 的文档。也就是 +printer +copier，但因为 + 在 URL 里面属于特殊字符，所以我们需要转义成 %2B，上面的 %20 代表空格的转义符。

如果你想查找只包含 A 不包含 B 的文档，可以 +A -B 来实现，这里我就不再演示了。关于 Solr 搜索的其他语法，这里就不再介绍了，感兴趣的人可以去官方文档里面阅读。

**本博客文章除特别声明，全部都是原创！**  
**原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。**  
**本文链接: 【】（）**