

solr接口说明文档

为了高效的进行搜索，使用搜索引擎代替SQL语句的搜索。

现在后端已经在服务器上部署了solr，并添加了中文分词器和相似匹配模块，实现了如下**两个接口**：

- 1. **科技成果的标题、摘要检索。**输入需要检索的内容，搜索引擎会进行分词等处理，返回需要的结果。

例如对于摘要进行检索，搜索“好的算法”，请求的url如下：

```
http://www.zdoubleleaves.cn:8080/solr/new_core/select?q=abstract:好的算法
```

会返回包含科技成果id的json：



The screenshot shows a web browser window with the address bar displaying the URL: `www.zdoubleleaves.cn:8080/solr/new_core/select?q=abstract:好的算法`. The browser's address bar also shows a security warning: "不安全" (Not Secure). Below the address bar, there are several search engine icons: 应用 (Applications), 北航校园网认证计... (Beihang University Network Authentication), 百度一下, 你就知道 (Baidu), Google, 大班公邮 (Daban Gongyou), QQ邮箱 (QQ Mail), and 百度翻 (Baidu Translate). The main content area of the browser displays a JSON response from the Solr search engine. The JSON is formatted with indentation and contains the following data:

```
{
  "responseHeader": {
    "status": 0,
    "QTime": 8,
    "params": {
      "q": "abstract:好的算法"
    }
  },
  "response": {
    "numFound": 2,
    "start": 0,
    "docs": [
      {
        "price": 0.0,
        "price_c": "0, USD",
        "id": "1",
        "abstract": "一种高效的算法",
        "title": ["PSO-KNN"],
        "url": "rsp/1.jpg",
        "_version_": 1633133421773455360,
        "price_c___l_ns": 0
      },
      {
        "price": 0.0,
        "price_c": "0, USD",
        "id": "12",
        "abstract": "一种也挺高校的算法",
        "title": ["算算算"],
        "url": "",
        "_version_": 1633133421796524032,
        "price_c___l_ns": 0
      }
    ]
  }
}
```

除了参数q，还可以使用别的参数，用于修改返回的内容等用途。前端同学可以学习一下参数的使用。

参数的使用参考博客：<https://blog.csdn.net/xiaocai9999/article/details/80385045>

- 2. **相似内容的推荐。**安装了MoreLikeThis组件，实现了对于相似内容进行打分排序，可以用于推荐相似的文章。

输入科技成果的id，可以返回与输入id的科技成果摘要内容相似内容的科技成果列表json。







例如，对于id=1的科技成果进行相似内容查询，想要返回相似内容的id和得分，请求的url如下：

```
http://www.zdoubleleaves.cn:8080/solr/new_core/select?
q=id:1&mlt=true&mlt.fl=abstract&mlt.mindf=1&mlt.mintf=1&fl=id,score
```

或

```
http://www.zdoubleleaves.cn:8080/solr/new_core/select?
q=id:1&mlt=true&mlt.fl=abstract&mlt.mindf=1&mlt.mintf=1&fl=id,score
```

← → ↻ ⓘ 不安全 | 123.206.68.192:8080/solr/new_core/select?q=id:1&mlt

 应用  北航校园网认证计...  百度一下，你就知道  Google  大班公邮 

```
{
  "responseHeader": {
    "status": 0,
    "QTime": 1,
    "params": {
      "q": "id:1",
      "mlt": "true",
      "fl": "id, score",
      "mlt.fl": "abstract",
      "mlt.mindf": "1",
      "mlt.mintf": "1"
    }
  },
  "response": {
    "numFound": 1,
    "start": 0,
    "maxScore": 2.1594841,
    "docs": [
      {
        "id": "1",
        "score": 2.1594841
      }
    ]
  },
  "moreLikeThis": {
    "1": {
      "numFound": 1,
      "start": 0,
      "maxScore": 2.3528168,
      "docs": [
        {
          "id": "12",
          "score": 2.3528168
        }
      ]
    }
  }
}
```

可以看到上图中的“moreLikeThis”的列表提供了id=12的科技成果。说明id=1的成果和id=12 的成果内容相似。

对象		achievement @test (123.20...	
开始事务		文本 · 筛选	排序 导入 导出
id ▾	title	abstract	time
1	PSO-KNN	一种高效的算法	2019-05
2	æµ<è ⁻ •æ ‡é¢~	æµ<è ⁻ •æ'~è	2019-05
3	æµ<è ⁻ •æ ‡é¢~	æµ<è ⁻ •æ'~è	2019-05
4	æµ<è ⁻ •æ ‡é¢~	æµ<è ⁻ •æ'~è	2019-05
5	æµ<è ⁻ •æ ‡é¢~	æµ<è ⁻ •æ'~è	2019-05
6	æµ<è ⁻ •æ ‡é¢~	æµ<è ⁻ •æ'~è	2019-05
7	æµ<è ⁻ •æ ‡é¢~	æµ<è ⁻ •æ'~è	2019-05
8	测试标题	测试摘要	2019-05
9	测试标题	测试摘要	2019-05
10	测试标题	测试摘要	2019-05
11	测试标题	测试摘要	2019-05
12	算算算	一种也挺高校的算法	2019-05

这是数据库科技成果表的内容，可以看到第1行和第12行的anstract列的内容较相似。

除了请求url的实例中的参数，mlt也有很多参数，用于修改搜索的要求和返回的内容。

参数的使用参考博客：<https://www.cnblogs.com/langfanyun/p/7490525.html> 或solrMLT文档