lucene：搜索引擎

说明：

|  |
| --- |
| Apache Lucene是一个开放源程序的搜寻器引擎，利用它可以轻易地为Java软件加入全文搜寻功能。Lucene的最主要工作是替文件的每一个字作索引，索引让搜寻的效率比传统的逐字比较大大提高，Lucen提供一组解读，过滤，分析文件，编排和使用索引的API，它的强大之处除了高效和简单外，是最重要的是使使用者可以随时应自己需要自订其功能。  　　Lucene是apache软件基金会[4] jakarta项目组的一个子项目，是一个开放源代码[5]的全文检索引擎工具包，即它不是一个完整的全文检索引擎，而是一个全文检索引擎的架构，提供了完整的查询引擎和索引引擎，部分文本分析引擎（英文与德文两种西方语言）。Lucene的目的是为软件开发人员提供一个简单易用的工具包，以方便的在目标系统中实现全文检索的功能，或者是以此为基础建立起完整的全文检索引擎。 |

相关度排序：优先级， 匹配度

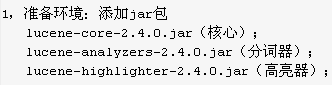
相关度高， 就排在最前面，

相关度算法。

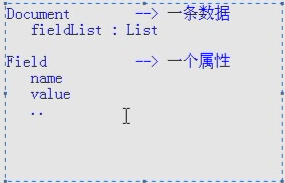
反向索引机制

： 根据词查文章。

|  |
| --- |
| 1. 建立索引 2. 从索引库中搜索 |



数据结构



要加入的jar包

|  |
| --- |
| lucen-core-3.0.1.jar(核心包)  contrib\analyzers\common\lucene-analyzers-3.0.1.jar(分词器)  contrib\highlighter\lucen-highlighter-3.0.1.jar(高亮)  comtrib\memory\lucene-memory-3.0.1.jar(lucen-highlighter-3.0.1.jar依赖它) |