搜索引擎开发指南

1.文档解析模块

- 使用Jsoup解析HTML

- 使用Tika解析PDF和Word

. 索引模块

- 使用Lucene建立索引

- 实现增量索引更新

. 搜索模块

- 实现多字段搜索

- 支持高亮显示