

数据库与信息处理

一种基于Lucene改进的全文检索工具包

宋佳^{1,2}, 诸云强¹, 刘润达^{1,2}

1.中国科学院 地理科学与资源研究所, 北京 100101

2.中国科学院 研究生院, 北京 100039

收稿日期 2007-5-25 修回日期 2007-5-25 网络版发布日期 2008-1-21 接受日期

摘要 基于Lucene实现了一种改进的全文检索引擎工具包ELucene。它引入了索引配置文件, 可针对不同应用背景来灵活定制索引的细节; 提供了定时自动更新索引的功能; 通过动态多态机制实现了支持多种索引数据源的功能; ELucene内部设计了引擎基础对象类, 并以静态对象的方式运行来避免频繁读取索引文件带来的性能损失。面向检索, 提供了检索请求类和检索响应类来分别封装用户的查询要求和查询结果集, 并设计实现了一些实用的查询输入和输出处理的方法。基于ELucene的元数据搜索系统已成功应用到“国家科学数据共享工程: 地球系统科学数据共享网”中。

关键词 [Lucene](#) [ELucene](#) [搜索引擎](#) [检索](#) [索引](#) [数据共享](#)

分类号

Enhanced full text retrieval kit based on Lucene

SONG Jia^{1,2}, ZHU Yun-qiang¹, LIU Run-da^{1,2}

1. Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

2. Graduate University of Chinese Academy of Sciences, Beijing 100039, China

Abstract

ELucene, an enhanced full text retrieval kit based on Lucene, is illustrated in this paper. The details of indexing can be customized via index configuration file according to different applications in ELucene. The function of periodic indices update is provided, and the support to kinds of data sources for indexing based on polymorphism is implemented. Class foundation objects in ELucene, which runs as static object mode, is designed for avoiding performance loss arising from frequent the disk indices reading. In searching process, the search request object and the search response object, in which some applied methods related to query input and query output are implemented, and provided for respectively encapsulating query requirements and query results. An application example of ELucene is given at the end of the paper.

Key words [Lucene](#) [ELucene](#) [search engine](#) [retrieval](#) [index](#) [data sharing](#)

DOI:

通讯作者 宋佳 songj@lreis.ac.cn

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(860KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [参考文献](#)

服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [复制索引](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

相关信息

▶ [本刊中包含“Lucene”的相关文章](#)

▶ [本文作者相关文章](#)

· [宋佳](#)

· [诸云强](#)

· [刘润达](#)