

<<XML文档全文检索的理论与方法>>

图书基本信息

书名：<<XML文档全文检索的理论与方法>>

13位ISBN编号：9787030319708

10位ISBN编号：7030319702

出版时间：2011-8

出版时间：科学

作者：夏立新|主编:王伟军

页数：171

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<XML文档全文检索的理论与方法>>

内容概要

本书是《知识管理与知识服务研究》丛书之一。

本书对XML全文检索的有关问题进行了全面系统的研究，既有对其基本理论、方法、技术的论述，也有对其最新研究进展的系统阐述。在深入分析文档的描述方法、用户的信息需求、检索方法和用户任务的基础上，总结了XML语言在文档描述、信息检索等方面的优势，研究了XML文档的索引实现方法，并以检索实验系统Okapi为例，探讨了如何在传统信息检索系统上实现XML文档的索引以支持XML检索并保持与原系统的兼容性，开发了基于XML的全文检索原型系统。

本书可作为计算机科学与技术、信息管理与信息系统、情报学、图书馆学等专业的教材或教学参考书，也可供信息中心、情报研究所、图书馆等机构的工作人员及广大信息用户学习参考。

作者简介

王伟军，华中师范大学教授、博士生导师。

现任华中师范大学信息管理系副主任；华中师范大学教学委员会委员、学位评定委员会管理学分会委员、知识管理与知识服务研究中心主任；兼任教育部高等学校图书馆学科教学指导委员会委员，中国索引学会常务理事，中国社会科学情报学会理事，中国信息经济学会电子商务专业委员会副秘书长，湖北省电子商务学会常务理事、副秘书长，湖北省信息学会常务理事；《情报科学》、《评价与管理》、《湖北信息化》等杂志的编委和《情报资料工作》学术指导委员会委员；曾担任第七届IFIP电子商务、电子服务与电子社会(I3E2007)国际会议主席。

2007年入选教育部新世纪优秀人才支持计划。

一直从事信息资源管理、知识管理与知识服务、电子商务等领域的教学和研究工作；公开发表学术论文80余篇，出版著作8部(含合著)；近5年来，主持或参与国家自然科学基金、国家社会科学基金、国家“863”计划、教育部高等学校学科创新引智计划、教育部人文社会科学研究和武汉市社会科学基金等10多项课题；曾荣获湖北省社会科学优秀成果一等奖和三等奖、湖北省自然科学优秀学术论文三等奖、第六届全国多媒体课件大赛优秀奖、武汉市社会科学优秀科研成果奖等多项奖励。

<<XML文档全文检索的理论与方法>>

书籍目录

总序

序

第1章 绪论

- 1.1 研究意义与研究背景
- 1.2 国内外研究现状
 - 1.2.1 XML上下文的相关性研究
 - 1.2.2 重复元素的过滤
 - 1.2.3 元素大小的处理
- 1.3 研究内容
- 1.4 结构安排

第2章 信息检索研究概述

- 2.1 信息检索的基本概念
 - 2.1.1 信息检索
 - 2.1.2 信息检索方法
 - 2.1.3 信息检索系统
 - 2.1.4 信息检索入口
- 2.2 文档描述与用户信息需求
 - 2.2.1 文档描述中的索引理论
 - 2.2.2 文档-索引词二维矩阵
 - 2.2.3 索引词的权重计算方法
 - 2.2.4 用户信息需求
- 2.3 信息检索方法
 - 2.3.1 布尔模型
 - 2.3.2 向量空间模型
 - 2.3.3 概率模型
- 2.4 信息检索系统
 - 2.4.1 信息检索系统构成
 - 2.4.2 信息检索原理
- 2.5 用户目标与任务
 - 2.5.1 用户目标
 - 2.5.2 用户目标与任务层次

第3章 XML置标语言

- 3.1 文档特征
 - 3.1.1 内容特征
 - 3.1.2 结构特征
 - 3.1.3 属性特征
 - 3.1.4 布局特征
- 3.2 文档置标与置标语言
 - 3.2.1 特殊置标和一般置标
 - 3.2.2 通用置标语言的发展
- 3.3 XML文档
 - 3.3.1 XML标准组成
 - 3.3.2 XML文档规范
 - 3.3.3 XML文档特征描述
 - 3.3.4 XML文档组织结构剖析

<<XML文档全文检索的理论与方法>>

- 3.4 良构的XML文档与XML名称空间
 - 3.4.1 XML语法规则
 - 3.4.2 建立一个良构的XML文档
 - 3.4.3 良构XML文档中的错误
 - 3.4.4 XML名称空间
- 3.5 XML文档的运行实例
 - 3.5.1 元素声明规则
 - 3.5.2 属性声明规则
 - 3.5.3 实体声明规则
 - 3.5.4 注释声明规则
- 3.6 XML文档的验证实例
 - 3.6.1 用DTD验证一个文档
 - 3.6.2 用XMLSchema描述一个文档
 - 3.6.3 RELAX NG技术
- 第4章 XML文档的索引方法
 - 4.1 基于XML文档的索引
 - 4.1.1 前序一后序节点标号法
 - 4.1.2 扩展的前序标号法
 - 4.1.3 基于路径的标号法
 - 4.1.4 节点定义法
 - 4.1.5 XML文档树的节点编码
 - 4.1.6 倒排索引
 - 4.1.7 结构索引
 - 4.1.8 索引策略
 - 4.2 基于XML文档的结构索引分析
 - 4.2.1 结构索引的方法及特点
 - 4.2.2 结构索引的潜在优势
 - 4.3 基于XML文档的索引机制的改良
- 第5章 XML信息检索
 - 5.1 XML信息检索的特点
 - 5.2 与XML相关的检索方式
 - 5.2.1 XML信息检索与XML数据库查询
 - 5.2.2 XML信息检索与HTML文档检索
 - 5.3 XML查询方法与查询语言
 - 5.3.1 XPath路径查询语言
 - 5.3.2 XQuery查询语言
 - 5.3.3 IR-Query, 一种扩展的XQuery查询语言
 - 5.4 XML检索查询代数
 - 5.4.1 传统关系代数表达式
 - 5.4.2 IAX查询代数
 - 5.4.3 XML查询操作符
 - 5.4.4 OrientXA查询代数
 - 5.5 面向XML信息检索的查询模型
 - 5.5.1 基于简单关键字的查询方式
 - 5.5.2 基于严格结构匹配与关键字相结合的查询模型
 - 5.5.3 基于非严格结构匹配与关键字相结合的查询模型
 - 5.6 XML信息检索的评价

<<XML文档全文检索的理论与方法>>

- 5.6.1 检索性能的评价
- 5.6.2 XML信息检索评价
- 5.6.3 INEX检索评价实验

第6章 基于传统文本检索系统的XML索引实现

- 6.1 传统文本检索系统概况
- 6.2 Okapi的索引结构
- 6.3 XML索引与实现
 - 6.3.1 XML文档预处理
 - 6.3.2 路径索引
 - 6.3.3 内容索引
- 6.4 实验评价
 - 6.4.1 XML数据集
 - 6.4.2 结果评价

第7章 基于XML的全文检索原型系统

- 7.1 基于中文XML文档的全文检索软件的原型系统开发
 - 7.1.1 系统设计目标
 - 7.1.2 全文检索系统的实现平台
 - 7.1.3 基于XML的全文检索系统的功能模块设计
 - 7.1.4 系统详细设计方案
 - 7.1.5 XML文档的解析
- 7.2 为XML文档建立全文索引
- 7.3 XML文档查询
- 7.4 基于XML的全文检索模型说明文档
 - 7.4.1 系统配置
 - 7.4.2 运行环境
 - 7.4.3 运行说明

第8章 结论

参考文献

附录

章节摘录

版权页：插图：信息检索入口又称检索点或检索标识，是指用以标识信息的外部特征和内部特征的属性值的集合。

检索标识是用户作为检索的出发点和依据，用户和信息检索系统之间的交流必须有一定的检索标识，否则会导致检索失败。

检索标识包括主题词、分类号、著者、标题、机构、代码等，它们通常由人工赋予，或由计算机自动生成。

2.2 文档描述与用户信息需求无论是非线性结构还是半结构文档，都是按一定结构组织的相关记录的集合。

本小节本着从用户信息需求的角度出发，对信息检索系统的逻辑视图等涉及文档描述的检索问题进行了分析，在索引的理论基础上，提出了文档一索引语词二维矩阵、文档一索引语词加权矩阵两种文档的描述方法，为XML文档的索引与检索问题的研究提供了理论和方法的支持。

<<XML文档全文检索的理论与方法>>

编辑推荐

《XML文档全文检索的理论与方法》为知识管理与知识服务研究之一。

<<XML文档全文检索的理论与方法>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>