

<<Internet主题搜索引擎设计与研究>>

图书基本信息

书名：<<Internet主题搜索引擎设计与研究>>

13位ISBN编号：9787508495781

10位ISBN编号：7508495780

出版时间：2012-3

出版时间：水利水电出版社

作者：梁春燕

页数：168

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>

## <<Internet主题搜索引擎设计与研究>>

### 内容概要

作为搜索引擎领域的重要发展趋势之一，主题搜索引擎充分考虑和满足用户对某些特定领域的网络信息需求，具有高度的目标化和专业化。

本书在继承国内外相关研究成果的基础上，以化学化工领域为例，设计了一个完整的主题搜索引擎系统，并基于专业知识库，研究了专业化爬行器、索引和检索的相关策略以及多语言自动分类技术和个性化技术，使Internet主题搜索引擎可提供更智能化、专业化和个性化的检索服务，更好地满足专业用户的信息需求。

本书内容丰富、应用性强，可供信息管理、计算机应用等领域从事相关研究的专家学者、工程技术人员及高等院校相关专业教师、研究生参考使用。

书籍目录

前言

常用符号表

第1章 Internet搜索引擎概述

- 1.1 Internet搜索引擎简介
- 1.2 Internet搜索引擎的基本原理
- 1.3 Internet搜索引擎的相关技术
  - 1.3.1 网络信息收集
  - 1.3.2 网络信息索引
  - 1.3.3 网络信息检索
  - 1.3.4 网页重要性评价技术-
  - 1.3.5 自动分类技术
  - 1.3.6 个性化技术
- 1.4 本书研究背景及内容
  - 1.4.1 本书研究背景
  - 1.4.2 本书研究内容

第2章 Internet主题搜索引擎的总体设计

- 2.1 Internet主题搜索引擎概述
- 2.2 Internet主题搜索引擎的目标
- 2.3 Internet主题搜索引擎的总体设计
- 2.4 Internet主题搜索引擎的基本模块
  - 2.4.1 爬行
  - 2.4.2 索引
  - 2.4.3 检索
  - 2.4.4 排序
  - 2.4.5 自动分类
  - 2.4.6 个性化检索
- 2.5 本章小结

第3章 Internet主题搜索引擎的信息收集和索引

- 3.1 Internet主题搜索引擎的收集策略
  - 3.1.1 爬行器的基本概念-
  - 3.1.2 爬行器的访问策略
  - 3.1.3 主题搜索引擎爬行器的设计和实现
- 3.2 Internet主题搜索引擎的索引策略
  - 3.2.1 索引器的基本概念
  - 3.2.2 主题搜索引擎索引器的设计和实现
- 3.3 本章小结

第4章 Internet主题搜索引擎的检索和排序

- 4.1 基于倒排索引的关键词检索
  - 4.1.1 检索策略
  - 4.1.2 检索的基本功能和实现
- 4.2 基于网络链接结构的网页评价和排序
  - 4.2.1 PageRank算法的实现
  - 4.2.2 基本排序方法
- 4.3 Internet主题搜索引擎的基本检索功能和相关界面
  - 4.3.1 基本检索

<<Internet主题搜索引擎设计与研究>>

4.3.2 使用偏好设置

4.3.3 高级检索

4.3.4 结果显示

4.4 本章小结

第5章 专业信息的自动分类

5.1 自动分类算法的研究

第6章 Internet主题搜索引擎的个性化检索

第7章 Internet主题搜索引擎的展望

参考文献

后记

章节摘录

版权页：插图：用户在使用搜索引擎进行信息检索时，通常会得到动辄几千几万的返回结果。面对这些数量巨大的信息，用户往往只会浏览前几页的结果，更多的结果则被埋在后面而不为用户所关心。

因此，将更能满足用户需求的文档优先显示给用户，就成为搜索引擎的重要目标，这通常是通过对搜索结果进行更好的排序来实现的。

在返回给用户的相关文档列表中，除了根据文档和用户查询的相关度进行排序外，还可以利用Internet上网页之间的相互评价，获取网页重要性/权威性的信息，参与到文档的排序中，从而可以将质量较好的文档优先显示给用户。

根据1.3.1节的阐述，可知网络上的资源通过超文本链接形成一个巨大的有向图（图1.4）。

从一个网页指向另一个网页的超链接，除了可以指导爬行器的爬行外，还可以看作是一个网页对链接网页的评价，这就构成了一个巨大的相互评价网络。

通过分析这个网络链接结构，使用某种算法来计算网页的链接重要性，使之与网页和查询的相关度一起共同决定返回网页的排序，就可以在同样相关度的情况下，将质量较好的网络资源优先排在前面。主要有两种网页重要性评价技术，即Google的PageRank算法和IBM的HITS技术。

编辑推荐

《Internet主题搜索引擎设计与研究》以化学化工专业为例，研究在专业知识库的基础上，自动收集和索引Internet专业资源并进行智能处理和智能检索的方法，研究建立Internet专业主题搜索引擎，为专业用户提供智能高效的网络检索服务。

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>