

<<现代信息检索>>

图书基本信息

书名：<<现代信息检索>>

13位ISBN编号：9787111331742

10位ISBN编号：7111331745

出版时间：2011-3

出版时间：机械工业出版社

作者：（西班牙）Ricardo Baeza-Yates,（巴西）Berthier Ribeiro-Neto

页数：913

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>

<<现代信息检索>>

内容概要

本书详细介绍了信息检索的所有主要概念和技术，以及有关信息检索方面的所有新变化，使读者既可以对现代信息检索有一个全面的了解，又可以获取现代信息检索所有关键主题的详细知识。

《现代信息融合技术在组合导航中的应用》的主要内容与信息检索领域的代表人物Baeza-Yares和JRibeiro-Neto编著，对于那些希望深入研究关键领域的读者，书中还提供了由其他主要研究人员编写的关于特殊主题的发展现状。

与上一版相比，本版在内容和结构上都有大量调整、更新和充实，其中新增内容在60%到70%左右。具体更新情况如下新增了文本分类、网络信息爬取、结构化文本检索和企业搜索等章节，以及关于开源搜索的一个附录。

·全面改写了用户界面、多媒体检索和数字图书馆等内容。

拓展了一些章节，介绍了信息检索方面的新的重要进展，如语言模型、新的评价方法、查询的特点、基于聚类和分布式信息检索等。

作者简介

Ricardo Baeza-Yates, 于加拿大滑铁卢大学获得计算机科学博士学位, 现为雅虎欧洲和拉丁美洲研究院副总裁, 主管雅虎在巴塞罗纳(西班牙)和圣地亚哥(智利)的研究中心, 并监管海法研究中心。

他曾担任智利计算机科学学会主席、智利大学计算机科学系Web研究中心主任、ICREA教授, 并且他还在巴塞罗纳法布拉大学创立了信息与通信技术系Web研究组。

现在他仍是智利大学和法布拉大学的兼职教授。

他的主要研究方向为算法与数据结构、信息检索、用户界面以及可视化在数据库中的应用等。

书籍目录

Preface to the Second Edition
 Preface to the First Edition
 Authors' Acknowledgements to the Second Edition
 Authors' Acknowledgements to the First Edition
 Publishers' Acknowledgements
 1 Introduction
 1.1 Information Retrieval
 1.1.1 Early Developments
 1.1.2 Information Retrieval in Libraries and Digital Libraries
 1.1.3 IR at the Center of the Stage
 1.2 The IR Problem
 1.2.1 The User's Task
 1.2.2 Information versus Data Retrieval
 1.3 The IR System
 1.3.1 Software Architecture of the IR System
 1.3.2 The Retrieval and Ranking Processes
 1.4 The Web
 1.4.1 A Brief History
 1.4.2 The e-Publishing Era
 1.4.3 How the Web Changed Search
 1.4.4 Practical Issues on the Web
 1.5 Organization of the Book
 1.5.1 Focus of the Book
 1.5.2 Book Contents
 1.6 The Book Web Site: A Teaching Resource
 1.7 Bibliographic Discussion
 User Interfaces for Search
 by Marti Hearst
 2.1 Introduction
 2.2 How People Search
 2.2.1 Information Lookup versus Exploratory Search
 2.2.2 Classic versus Dynamic Model of Information Seeking
 2.2.3 Navigation versus Search
 2.2.4 Observations of the Search Process
 2.3 Search Interfaces Today
 2.3.1 Getting Started
 2.3.2 Query Specification
 2.3.3 Query Specification Interfaces
 2.3.4 Retrieval Results Display
 2.3.5 Query Reformulation
 2.3.6 Organizing Search Results
 2.4 Visualization in Search Interfaces
 2.4.1 Visualizing Boolean Syntax
 2.4.2 Visualizing Query Terms within Retrieval Results
 2.4.3 Visualizing Relationships Among Words and Documents
 2.4.4 Visualization for Text Mining
 2.5 Design and Evaluation of Search Interfaces
 2.6 Trends and Research Issues
 2.7 Bibliographic Discussion
 Modeling
 3.1 IR Models
 3.1.1 Modeling and Ranking
 3.1.2 Characterization of an IR Model
 3.1.3 A Taxonomy of IR Models
 3.2 Classic Information Retrieval
 3.2.1 Basic Concepts
 3.2.2 The Boolean Model
 3.2.3 Term Weighting
 3.2.4 TF-IDF Weights
 3.2.5 Document Length Normalization
 3.2.6 The Vector Model
 3.2.7 The Probabilistic Model
 3.2.8 Brief Comparison of Classic Models
 3.3 Alternative Set Theoretic Models
 3.3.1 Set-Based Model
 3.3.2 Extended Boolean Model
 3.3.3 Fuzzy Set Model
 3.4 Alternative Algebraic Models
 3.4.1 Generalized Vector Space Model
 3.4.2 Latent Semantic Indexing Model
 3.4.3 Neural Network Model
 3.5 Alternative Probabilistic Models
 3.5.1 BM25
 3.5.2 Language Models
 3.5.3 Divergence from Randomness
 3.5.4 Bayesian Network Models
 3.6 Other Models
 4 Retrieval Evaluation
 5 Relevance Feedback and Query Expansion
 6 Documents: Languages & Properties
 7 Queries: Languages & Properties
 8 Text Classification
 9 Indexing and Searching
 10 Parallel and Distributed IR
 11 Web Retrieval
 12 Web Crawling
 13 Structured Text Retrieval
 14 Multimedia Information retrieval
 15 Enterprise Search
 16 Library Systems
 17 Digital Libraries

章节摘录

Libraries were among the first institutions to adopt IR systems for retrieving information. Usually, library systems were initially developed by academic institutions and later by commercial vendors. In the first generation, such systems consisted of an automation of existing processes such as card catalogs searching, restricted to author names and titles. In the second generation, increased search functionality was added to include subject headings, keywords, and query operators. In the third generation, which is currently being deployed, the focus has been on improved graphical interfaces, electronic forms, hypertext features, and open system architectures.

Traditional library management system vendors include Endeavor Information Systems Inc., Innovative Interfaces Inc., and EOS International. Among systems developed with a research focus, we distinguish MELVYL developed by the California Digital Library at University of California, and the Cheshire system developed originally at UC Berkeley and lately in cooperation with the University of Liverpool. Further details on these library systems can be found in Chapter 16.1.1.3 IR at the Center of the Stage. Despite its maturity, until recently, IR was seen as a narrow area of interest restricted mainly to librarians and information experts. Such a tendentious vision prevailed for many years, despite the rapid dissemination, among users of modern personal computers, of IR tools for multimedia and hypertext applications. In the beginning of the 1990s, a single fact changed once and for all these perceptions: the introduction of the World Wide Web. The Web, invented in 1989 by Tim Berners-Lee, has become a universal repository of human knowledge and culture. Its success is based on the conception of a standard user interface which is always the same, no matter the computational environment used to run the interface, and which allows any user to create their own documents. As a result, millions of users have created billions of documents that compose the largest human repository of knowledge in history. An immediate consequence is that finding useful information on the Web is not always a simple task and usually requires posing a query to a search engine, i.e., running a search. And search is all about IR and its technologies. Thus, almost overnight, IR has gained a place with other technologies at the center of the stage. ……

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>