信息检索导论（修订版）

[信息检索导论（修订版）_下载链接1_]

著者:普拉巴卡尔·拉格万（Prabhakar Raghavan）

出版者:人民邮电出版社

出版时间:2019-7-1

装帧:平装

isbn:9787115514080

本书是信息检索的教材，旨在从计算机科学的视角提供一种现代的信息检索方法。书中从基本概念讲解网络搜索以及文本分类和文本聚类等，对收集、索引和搜索文档系统的设计和实现的方方面面、评估系统的方法、机器学习方法在文本收集中的应用等给出了

最新的讲解。

什么是排序SVM、XML、DNS和LSI？什么是信息检索中的垃圾信息、隐藏页和门页？MapReduce和其他一些并行运算方法是如何实现由兆字节到百万兆字节的飞跃的？这些问题你都能从本书中找到答案。本书首次将构建Web搜索引擎的复杂过程以一种清晰的全景方式展现给读者。——Peter Norvig，计算机科学家，Google研发总监

本书对信息检索这个举足轻重、发展迅猛的领域进行了全面、准确的介绍，是一本不可多得的教材。——Raymond Mooney，得克萨斯大学奥斯汀分校教授

本书选材独特，对信息检索的基础知识和发展方向进行了生动描述。——Jon Kleinberg，康奈尔大学教授

作者介绍:

【美】克里斯托夫·曼宁 （Christopher Manning）
计算机科学家，斯坦福大学教授，斯坦福大学人工智能实验室主任，ACM会士、AAAI会士、ACL会士。目前的研究目标为计算机如何智能地处理、理解和生成人类语言资料。曼宁博士是深度学习在自然语言处理应用方面的先锋人物，在树递归神经网络、语义分析、神经机器翻译、深度语言理解等方面均有令业界瞩目的研究成果。

【美】普拉巴卡尔·拉格万（Prabhakar Raghavan）Google高级副总裁，目前负责谷歌的广告与商业产品、基础设施团队。之前作为Google App和Google Cloud的副总裁，带领团队做出了突出业绩。在加入Google前任职于Yahoo!，是Yahoo!实验室的创建者和负责人。拉格万博士毕业于加州大学伯克利分校，长期担任斯坦福大学计算机科学系顾问教授，主要研究方向是文本及Web数据挖掘、随机算法等，是美国国家工程院院士、ACM会士、IEEE会士。

【德】欣里希·舒策（Hinrich Schütze）
德国慕尼黑大学信息与语言处理中心主任，计算语言学家，斯坦福大学博士。曾在美国硅谷工作多年。

王斌
博士，小米公司AI实验室NLP方向首席科学家，前中国科学院信息工程研究所研究员、博导，中国科学院大学教授。

李鹏 博士，中国科学院信息工程研究所高级工程师，硕士生导师。

目录: 第 1章 布尔检索 ． 1

・・・・・・・(收起)

信息检索导论（修订版）_下载链接1_

标签

搜索引擎

计算机科学

数据库

CS

评论

对于原理讲得很清晰，逐步引导读者实现搜索引擎，加深对信息检索的理解

作为导论可以打100分，其中有一些关于语音搜索的点非常有启发性，对业务价值很大。Mark：https://libindic.org/Soundex

---
信息检索导论（修订版）_下载链接1_

书评

对于搜索引擎的初学者里说，本书是一本绝对值得阅读的书目。作者从最简单的布尔检索到一个完整的搜索引擎，逐步深入，逐步引导读者思考，对建造一个大型搜索引擎需要用到的架构和算法都有所涉猎，看完后会对搜索引擎有一个大概的认识，对其基本原理也会有所了解。搜索引擎并不...

---
stanford的IR入门书籍，cmu
stanford都在用该书作为IR入门书籍，很nice。在某些章节如果你有统计的基础来看的话，会更容易些。

第一次看到这本书的时候，还是在前年，当时这本书还只是个草稿的电子版，基本上ir所涉及到的内容都有，讲的也比较全面。
要是你英文阅读能力还好的话，推荐去读读这本书，肯定会对ir有一个较为全面的了解的。

最重要的收获，是对信息检索系统（搜索引擎）有一个宏观的认识，大体上说，需要从两个维度来看：
第一个是查询维度，它的核心，是两个索引结构；其一是字典，其二是倒排拉链和正排索引； 字典的职责，是把 query 变成 term
set；期间用到了多种技术，如：语义扩展（同义词、拼...

这本书不错。值得一看。 Christopher D.
Manning，1989年毕业于澳大利亚国立大学，1995年获斯坦福大学语言学博士学位，曾先后在卡内基-梅隆大学、悉尼大学教授语言学，1999年起任斯坦福大学计算机科学和语言学副教授，其主要研究方向是统计自然语言处理、信息提取与表示，以及...

作为入门书籍，还不错。分别介绍了信息检索领域的几个重要概念：倒排索引、检索引擎；tf-idf权重计算技术；向量空间模型，信息检索的评价，有序检索结果的评价MAP，ROC曲线，NDCG等等；相关反馈技术，伪相关反馈；概率检索模型，BM25算法；基于语言建模的信息检索模型，各种文...

搜素引擎入门书籍，各方面均有涉猎，严谨，通俗易懂
入门经典

信息检索导论（修订版）_下载链接1_