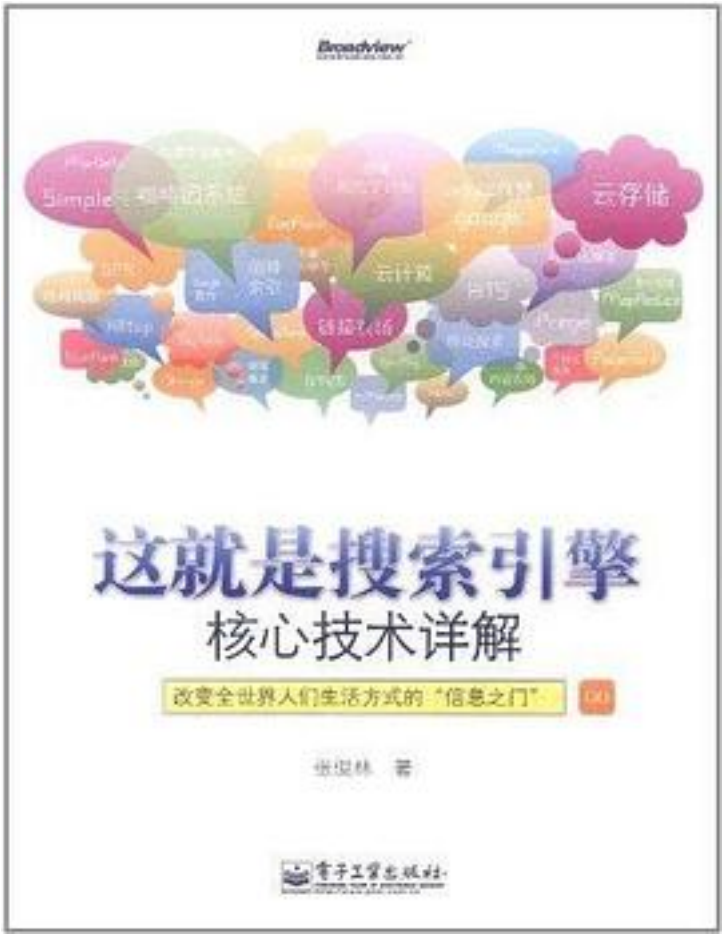
这就是搜索引擎



这就是搜索引擎_下载链接1_

著者:张俊林

出版者:电子工业出版社

出版时间:2012-1-1

装帧:平装

isbn:9787121148651

搜索引擎作为互联网发展中至关重要的一种应用，已经成为互联网各个领域的制高点，其重要性不言而喻。搜索引擎领域也是互联网应用中不多见的以核心技术作为其命脉的领域，搜索引擎各个子系统是如何设计的？这成为广大技术人员和搜索引擎优化人员密

切关注的内容。

本书的最大特点是内容新颖全面而又通俗易懂。对于实际搜索引擎所涉及的各种核心技术都有全面细致的介绍，除了作为搜索系统核心的网络爬虫、索引系统、排序系统、链接分析及用户分析外，还包括网页反作弊、缓存管理、网页去重技术等实际搜索引擎必须关注的技术，同时用相当大的篇幅讲解了云计算与云存储的核心技术原理。另外，本书也密切关注搜索引擎发展的前沿技术：Google的咖啡因系统及Megastore等云计算新技术、百度的暗网抓取技术阿拉丁计划、内容农场作弊、机器学习排序等。诸多新技术在相关章节都有详细讲解，同时对于社会化搜索、实时搜索及情境搜索等搜索引擎的未来发展方向做了技术展望。为了增进读者的理解，全书大量引入形象的图片来讲解算法原理，相信读者会发现原来搜索引擎的核心技术理解起来比原先想象的要简单得多。

作者介绍:

张俊林：本科毕业于天津大学管理学院，2004年于中科院软件所直接获得博士学位并留所从事科研工作，研究方向为搜索引擎与自然语言处理。2005年在CSDN博客发布系列博文“搜索引擎设计实用教程：以百度为例”，在网络上获得了广泛转载与良好口碑。2006年作为联合创始人建立了智能信息聚合网站“玩聚网”，曾先后于阿里巴巴搜索技术中心任资深搜索技术研究员、房价网首席研究员，现任职于新浪微博，从事微博搜索与语义分析及推荐方面的研发工作。

目录: 目 录

这就是搜索引擎_下载链接1_

标签

搜索引擎

信息检索

互联网

数据挖掘

算法

搜索

计算机

评论

逻辑清晰,作者很用心；火车上看的,三五六八是重点。

对咱们本行来说没有价值, 该知道的都知道了. 但对入门的人来说是极品.

挺好的入门书

先吐槽，不论装帧，书里边的图例和各种修饰图案就非常不适合出现在一本技术书当中，儒家思想都出来了，作者写好内容就可以了，其实大可不必往书里面加一些这么标新立异的东西，全是违和感好么。说完图就说排版，比word还难看，给文字内容造成的负面影响极大。不论内容，这就是本糟糕极了的书，可以作为出版界的反面教材。如果单体内容，怎么说呢我虽然才开始看相关的内容，但我觉得应该还是有干货的，即便二手的很多。把云存储单独成章我觉得有点突兀，介绍的时候和信息检索的联系又很少。好了负面评价这么多，那我还是给4星是因为好多人都给4星，技术方面又是新手，不敢妄自菲薄。不过还是支持吧。

: G254.4/1224

虽然是对各种技术的摘抄，但摘的不错，阅读起来很流畅，对于初学者来说，比起花大量时间到处查阅各种资料，花一天时间集中精力看完本书还是很划算的

非码农跳着看也能看懂，为数不多的注意英文前后空格的中文书，可惜估计找错出版社了……制作之糙，结尾都没一个。

就像是写综述，别想通过此书学会什么，只能作为一个各种技术的索引，知道有哪些研究领域。对我个人而言比较新颖的就是网页去重，原来没意识到这块也有很多工作。

图例丰富；隔靴搔痒

有点过时了，微信读书的排版很极品，但貌似目前也没有很好的搜索引擎中文科普了。

书丢飞机上了。 看过大半

1

出乎意料的好

连科普都看不进去。。

一本入门书，涵盖工程及算法。

讲述太啰嗦了，不够精炼。

比较通俗，可惜对细节涉及不深，每章最后自带paper references

入门书都给人一种什么东西都是故弄玄虚其实秒懂的错觉

其实我是用这本书来看有关随机过程的应用滴~~嘿嘿，而且我跟作者在微博上互动很多。。

科普读物，对于我这种门外汉最合适。

这就是搜索引擎_下载链接1_

书评

个人一直都对搜索引擎有着比较高的兴趣，算上这本书，应该一共看过三本搜索引擎相关的书： 1. 深入搜索引擎 : 海量信息的压缩、索引和查询 2. 自己动手写搜索引擎 3. 本书
第二本书无须多说，我觉得自己动手系列简直就是一坨屎一样的存在，而第一本书非常详细地讲了关于搜索...

正如作者前言中提到，搜索引擎技术培训，发现缺少相关的入门书，
要么太理论，要么太深入，初学者的门槛很高，其实对于初学者来说，最想快速了解到的，这是一个什么东西，做什么，大体流程是什么样，总体有了认识之后，再逐步挑自己喜欢的地方深入的学习和研究，到这时候，那些...

整本书的风格是广而不深，对搜索引擎相关的很多技术都进行了介绍，每个方面的介绍基本都在一页纸的篇幅，同时在每章后面附带了若干参考文献以供求职欲强的读者查询。
作者基本做到了浅出，全书没有代码，基本也没有什么数学公式，全部用图来展现，读者即使没有相关基础...

最近几年一直坚持一个观点，那就是IT类尤其是IT技术类的书籍，只看英文原文的。原文还尽量找非扫描的PDF格式的，为的是可以尽情的标注、记笔记和搜索。后来还弄来一部9寸的kindle dxg。 这本书改变了我的看法，行文流畅，插图生动，深入浅出。
一周零零散散的读完，一句话概括...

这本书花了我将近四个晚上的时间读完。
全书内容通俗易懂，读的过程中非常愉快，当合上这本书的时候从里到外有一种很舒畅的感觉。上次读技术书籍有这种感觉，我记得是在读Head First系列书籍的时候。
此书在内容的讲解上很清晰，图文并茂。将算法和原理解释的清晰易懂...

基本覆盖到了大规模商业搜索引擎的各个关键技术点（基本还是以全网搜索为背景的），花了一个周末读完，感觉比较大的帮助是把之前自己所了解的零散的技术点组织成体系，不过瘾的是对于单个技术点并没有深入的阐述。
总而言之，是一本很好的商业搜索引擎研发入门书籍（之前好像没...

刚通读读完这本书，受益匪浅啊，对搜索技术总算有一个体系上的认知了，准备再精读一遍，通俗易懂
，很复杂的问题能表述的简单明了，好的技术书应该读起来轻盈而优雅，正如本书。而不是充满苦涩难懂，，强烈推荐给对搜索引擎有兴趣的程序员！！！！！！！！！！！！！！！！！！...

第20页 2.4.1 宽度优先遍历策略（Breath First）应该为Breadth First,既然不是翻译的，是我们中国人自己写的书，更应该严谨。、、
...

全文搜索引擎，涉及面很广，很全。但给人感觉具体点上比较简单。结构方面，整体脉络比较清晰，看过之后，给人很多启发。实际使用中被忽略的一些点，在书中可以找到很多参考的东西。大数据时代的章节内容比较详实，存储系统框架介绍不错。再结合其他搜索书籍看更好。 希望后续可...

利用周末及工作下班之余, 每天30分钟~1小时, 断断续续地搞了3周.
豆瓣上对这本书的评价还是不错的, 但是对于一名电商领域的工程人员来说,
其实用价值却不太大,
毕竟网络爬虫、检索模型、链接分析、查询意图分析这些东东在平时工作中很少接触到, 但对于知识面的拓广、眼界的开...

工作上开始接触到搜索引擎的内容，为了能够快速入门，选了这本书。粗读了一遍，确实感觉到把搜索引擎的内容做了比较全面和浅显的了解，里面讲解的算法都是比较浅显

的说明了一下。
唯一的遗憾是我比较关心的与地理位置相关的搜索没有详细的说明，只在最后一章当做展望...

首先要说，印刷质量不错，纸张很厚。
此书通俗易懂，插图也比较多，不过插图也起不到画龙点睛的效果，好像就是为了插图而插图。跟老外的比起来，写作方法还欠缺。
可能是此书就是科普性质的，广而不精，可以晚上翻看，甚至在公交车上看，作为扩大知识面或者说入门还是很有...

此书总体上还可以，在讲一个算法是什么的时候下了挺大功夫，例如画了很多图，基本每个算法都举了简单的例子，这点还是不错的，值得鼓励。但是像很多国内的技术书一样，在讲why的时候不怎么给力，当然这个要求比较高，总体上还是值得推荐

除了索引压缩那一章，都看了。
不过发现当时看懂了，看后边章节的时候有的前边的就忘了
想不起细节，还是要和实践结合
学知识最好的方法就是去用知识，不然看了不久就忘光了
.---------- ----------...

这就是搜索引擎_下载链接1_