Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems)

By Ian H. Witten, Alistair Moffat, Timothy C. Bell

Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell

In this fully updated second edition of the highly acclaimed **Managing Gigabytes**, authors Witten, Moffat, and Bell continue to provide unparalleled coverage of state-of-the-art techniques for compressing and indexing data. Whatever your field, if you work with large quantities of information, this book is essential reading--an authoritative theoretical resource and a practical guide to meeting the toughest storage and access challenges. It covers the latest developments in compression and indexing and their application on the Web and in digital libraries. It also details dozens of powerful techniques supported by mg, the authors' own system for compressing, storing, and retrieving text, images, and textual images. mg's source code is freely available on the Web.

- Up-to-date coverage of new text compression algorithms such as block sorting, approximate arithmetic coding, and fat Huffman coding
- New sections on content-based index compression and distributed querying, with 2 new data structures for fast indexing
- New coverage of image coding, including descriptions of de facto standards in use on the Web (GIF and PNG), information on CALIC, the new proposed JPEG Lossless standard, and JBIG2
- New information on the Internet and WWW, digital libraries, web search engines, and agent-based retrieval
- Accompanied by a public domain system called MG which is a fully worked-out operational example of the advanced techniques developed and explained in the book
- New appendix on an existing digital library system that uses the MG software

⬇ **Download** Managing Gigabytes: Compressing and Indexing Docum

...pdf

Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems)

By Ian H. Witten, Alistair Moffat, Timothy C. Bell

Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell

In this fully updated second edition of the highly acclaimed **Managing Gigabytes**, authors Witten, Moffat, and Bell continue to provide unparalleled coverage of state-of-the-art techniques for compressing and indexing data. Whatever your field, if you work with large quantities of information, this book is essential reading--an authoritative theoretical resource and a practical guide to meeting the toughest storage and access challenges. It covers the latest developments in compression and indexing and their application on the Web and in digital libraries. It also details dozens of powerful techniques supported by mg, the authors' own system for compressing, storing, and retrieving text, images, and textual images. mg's source code is freely available on the Web.

- Up-to-date coverage of new text compression algorithms such as block sorting, approximate arithmetic coding, and fat Huffman coding
- New sections on content-based index compression and distributed querying, with 2 new data structures for fast indexing
- New coverage of image coding, including descriptions of de facto standards in use on the Web (GIF and PNG), information on CALIC, the new proposed JPEG Lossless standard, and JBIG2
- New information on the Internet and WWW, digital libraries, web search engines, and agent-based retrieval
- Accompanied by a public domain system called MG which is a fully worked-out operational example of the advanced techniques developed and explained in the book
- New appendix on an existing digital library system that uses the MG software

Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell Bibliography

- Sales Rank: #893643 in Books
- Brand: Brand: Morgan Kaufmann
- Published on: 1999-05-17
- Original language: English
- Number of items: 1
- Dimensions: 9.54" h x 1.34" w x 7.67" l, 2.49 pounds
- Binding: Hardcover
- 550 pages

Download and Read Free Online Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell

Editorial Review

Amazon.com Review
Of all the tasks programmers are asked to perform, storing, compressing, and retrieving information are some of the most challenging--and critical to many applications. *Managing Gigabytes: Compressing and Indexing Documents and Images* is a treasure trove of theory, practical illustration, and general discussion in this fascinating technical subject.

Ian Witten, Alistair Moffat, and Timothy Bell have updated their original work with this even more impressive second edition. This version adds recent techniques such as block-sorting, new indexing techniques, new lossless compression strategies, and many other elements to the mix. In short, this work is a comprehensive summary of text and image compression, indexing, and querying techniques. The history of relevant algorithm development is woven well with a practical discussion of challenges, pitfalls, and specific solutions.

This title is a textbook-style exposition on the topic, with its information organized very clearly into topics such as compression, indexing, and so forth. In addition to diagrams and example text transformations, the authors use "pseudo-code" to present algorithms in a language-independent manner wherever possible. They also supplement the reading with *mg*--their own implementation of the techniques. The mg C language source code is freely available on the Web.

Alone, this book is an impressive collection of information. Nevertheless, the authors list numerous titles for further reading in selected topics. Whether you're in the midst of application development and need solutions fast or are merely curious about how top-notch information management is done, this hardcover is an excellent investment. *--Stephen W. Plain*

Topics covered: Text compression models, including Huffman, LZW, and their variants; trends in information management; index creation and compression; image compression; performance issues; and overall system implementation.

Review
"This book is the Bible for anyone who needs to manage large data collections. It's required reading for our search gurus at Infoseek. The authors have done an outstanding job of incorporating and describing the most significant new research in information retrieval over the past five years into this second edition."-**Steve Kirsch, Cofounder, Infoseek Corporation**

"The new edition of Witten, Moffat, and Bell not only has newer and better text search algorithms but much material on image analysis and joint image/text processing. If you care about search engines, you need this book: it is the only one with full details of how they work. The book is both detailed and enjoyable; the authors have combined elegant writing with top-grade programming."-**Michael Lesk, National Science Foundation**

"The coverage of compression, file organizations, and indexing techniques for full text and document management systems is unsurpassed. Students, researchers, and practitioners will all benefit from reading this book."-**Bruce Croft, Director, Center for Intelligent Information Retrieval at the University of**

Massachusetts

From the Back Cover

"This book is the Bible for anyone who needs to manage large data collections. It's required reading for our search gurus at Infoseek. The authors have done an outstanding job of incorporating and describing the most significant new research in information retrieval over the past five years into this second edition."
Steve Kirsch, Cofounder, Infoseek Corporation

"The new edition of Witten, Moffat, and Bell not only has newer and better text search algorithms but much material on image analysis and joint image/text processing. If you care about search engines, you need this book: it is the only one with full details of how they work. The book is both detailed and enjoyable; the authors have combined elegant writing with top-grade programming."
Michael Lesk, National Science Foundation

"The coverage of compression, file organizations, and indexing techniques for full text and document management systems is unsurpassed. Students, researchers, and practitioners will all benefit from reading this book."
Bruce Croft, Director, Center for Intelligent Information Retrieval at the University of Massachusetts

In this fully updated second edition of the highly acclaimed **Managing Gigabytes**, authors Witten, Moffat, and Bell continue to provide unparalleled coverage of state-of-the-art techniques for compressing and indexing data. Whatever your field, if you work with large quantities of information, this book is essential reading--an authoritative theoretical resource and a practical guide to meeting the toughest storage and access challenges. It covers the latest developments in compression and indexing and their application on the Web and in digital libraries. It also details dozens of powerful techniques supported by mg, the authors' own system for compressing, storing, and retrieving text, images, and textual images. mg's source code is freely available on the Web.

Users Review

From reader reviews:

Emma Patterson:

Hey guys, do you wants to finds a new book you just read? May be the book with the subject Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) suitable to you? Typically the book was written by famous writer in this era. The particular book untitled Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems)is the main one of several books that everyone read now. This particular book was inspired lots of people in the world. When you read this book you will enter the new dimensions that you ever know just before. The author explained their plan in the simple way, and so all of people can easily to know the core of this reserve. This book will give you a lot of information about this world now. To help you see the represented of the world in this particular book.

Augustus Chase:

Reading a e-book can be one of a lot of pastime that everyone in the world likes. Do you like reading book so. There are a lot of reasons why people enjoy it. First reading a book will give you a lot of new facts. When you read a book you will get new information because book is one of various ways to share the information as well as their idea. Second, studying a book will make you more imaginative. When you examining a book especially fiction book the author will bring one to imagine the story how the figures do it anything. Third, you are able to share your knowledge to other folks. When you read this Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems), you may tells your family, friends in addition to soon about yours e-book. Your knowledge can inspire others, make them reading a book.

Kathleen Hernandez:

Do you have something that you enjoy such as book? The book lovers usually prefer to select book like comic, limited story and the biggest one is novel. Now, why not seeking Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) that give your satisfaction preference will be satisfied through reading this book. Reading habit all over the world can be said as the method for people to know world a great deal better then how they react when it comes to the world. It can't be stated constantly that reading habit only for the geeky individual but for all of you who wants to possibly be success person. So , for all you who want to start reading through as your good habit, you could pick Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) become your current starter.

Donald Purcell:

You will get this Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) by browse the bookstore or Mall. Only viewing or reviewing it might to be your solve issue if you get difficulties on your knowledge. Kinds of this publication are various. Not only by written or printed but in addition can you enjoy this book simply by e-book. In the modern era just like now, you just looking of your mobile phone and searching what your problem. Right now, choose your own personal ways to get more information about your guide. It is most important to arrange yourself to make your knowledge are still revise. Let's try to choose right ways for you.

Download and Read Online Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell #2ATEGCQ6X1S

Read Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell for online ebook

Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell books to read online.

Online Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell ebook PDF download

Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell Doc

Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell Mobipocket

Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell EPub

2ATEGCQ6X1S: Managing Gigabytes: Compressing and Indexing Documents and Images, Second Edition (The Morgan Kaufmann Series in Multimedia Information and Systems) By Ian H. Witten, Alistair Moffat, Timothy C. Bell