MORGAN KAUFMANN PUBLISHERS, INC.

MANAGING GIGABYTES

COMPRESSING AND INDEXING DOCUMENTS AND IMAGES

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

Second edition; Spring 1999; 500 pages; cloth; 1-55860-570-3; $59.95

As today's information explosion generates greater and greater volumes of raw data, the challenge of storing and retrieving this information in the most efficient manner continues to grow, whether the data is stored on a local disk or distributed over the World-Wide Web.

Managing Gigabytes helps you to meet this challenge by showing how to capitalize on new methods of compressing and accessing data, enabling you to store information more efficiently and locate specific items more quickly and cost-effectively than ever before. It uniquely covers fully-tested techniques for both text and image compression and shows how to construct a tailor-made electronic index for accessing text, scanned documents, and images.

This book largely avoids extensive theoretical and mathematical discussions, making it accessible to curious laypersons who seek a clear, uncomplicated understanding of this new technology. Real, large-scale problems are illustrated, and the technical material is sprinkled with anecdotes and background information. The new edition is updated with information about recent standards and discoveries. An instructor's supplement will be available.

Managing Gigabytes provides current and comprehensive tools and techniques that will help professionals and academics to work more confidently and effectively in today's increasingly paperless society.

NEW FEATURES

All chapters have been thoroughly updated and new references added to the recent literature;

Up-to-the-minute information on new text compression algorithms is included such as block sorting, approximate arithmetic coding, and fast Huffman coding;

New sections on context-based index compression and on distributed querying, and two new data structures for fast indexing;

Significant new material on image coding like descriptions of de facto standards that are in widespread use on the Web (GIF and PNG), information on CALIC, a high-performance lossless scheme, the new proposed "JPEG Lossless" international standard, and the new JBIG2 international standard for compressing document images;

New information on the Internet and World-Wide Web, digital libraries, Web search engines, and agent-based information retrieval;

New Appendix describing the New Zealand Digital Library, a large-scale web-based application incorporating many of the key ideas presented in the book;

An Instructor's Supplement will be available that contains review and test questions.

TABLE OF CONTENTS

ABOUT THE AUTHORS

IAN H. WITTEN is Professor of Computer Science at the University of Waikato in New Zealand. He is a Fellow of the ACM and of the Royal Society of New Zealand, and a member of professional computing, information retrieval, and engineering associations in the UK, USA, Canada, and New Zealand. He is co-author of *The Reactive Keyboard* (1992) and *Text Compression* (1990), as well as many journal articles and conference papers.

ALISTAIR MOFFAT is Associate Professor of Computer Science at the University of Melbourne. He is the author of numerous peer-reviewed papers, which have explored such areas as algorithms and data structures for text and image compression, self-adjusting data structures for dictionaries and priority queues, and algorithms for adaptive searching and sorting.

TIMOTHY C. BELL is a Senior Lecturer in Computer Science at the University of Canterbury. He is co-author of *Text Compression* (1990) and the author of a number of papers covering text and image compression, as well as computers and music, and computer education.

Order from Morgan Kaufmann Publishers

Mail: Harcourt Brace & Company Australia, Locked Bag 16
Marrickville, NSW 2204, Australia
Phone: 61-2-9517-8999, **Fax:** 61-2-9517-2249

Mail: Harcourt Brace & Co., Attn. Order Fulfillment Dept.,
6277 Sea Harbor Drive, Orlando, FL 32887,
Phone: 800-745-7323 , 407-345-3800, **Fax:** 800-874-6418, 407-345-4060,
Email: orders@mkp.com, **Visit Morgan Kaufmann on the Web:** www.mkp.com