
Apache Mahout

Beyond Mapreduce

Big Data and Hadoop
Field Guide to Hadoop
Professional Hadoop Solutions
Handbook of Big Data Technologies
Data-intensive Text Processing with MapReduce
Intelligence in Big Data Technologies—Beyond
the Hype
Big Data, Big Challenges: A Healthcare
Perspective
Hadoop: The Definitive Guide
Artificial Intelligence: Concepts, Methodologies,
Tools, and Applications
Big Data Using Hadoop and Hive
Apache Mahout Essentials
Apache Mahout
Database Systems for Advanced Applications
Hadoop 2 Quick-Start Guide
Machine Learning and the Internet of Medical
Things in Healthcare
The Web at Graduation and Beyond
Software Reuse for Dynamic Systems in the Cloud
and Beyond
Big Data Analytics Beyond Hadoop
Future And Fintech, The: Abcdi And Beyond
Intelligent Systems
Machine Learning
Hadoop in Action

Resilience in the Digital Age
Learning from Imbalanced Data Sets
Big Data at Work
Analytics and Big Data: The Davenport Collection
(6 Items)
Contemporary Research on E-business
Technology and Strategy
Kubeflow for Machine Learning
Cloud Architecture Patterns
Cloud Computing
Handbook of Research on Engineering Education
in a Global Context
Mahout in Action
Hadoop: The Definitive Guide
Hadoop in Practice
Big Data Analytics
Handbook of E-Tourism
Enterprise Data Workflows with Cascading
Data Communication and Networks
Professional NoSQL

Apache Mahout Beyond Mapreduce Downloaded from db.mwpai.edu by guest

BEST HARVEY

Big Data and Hadoop BoD - Books on Demand
The Future and FinTech

examines the fundamental financial technologies and its growing impact on the Banking, Financial Services and Insurance (BFSI) sectors. With global investment amounting to more than \$100 billion in 2020, the proliferation of FinTech has underpinned the direction

payments, loans, wealth management, insurance, and cryptocurrencies are heading. This book presents FinTech from an industrial perspective in the context of architecture and its basic building blocks, e.g., Artificial Intelligence (AI), Blockchain, Cloud, Big Data, Internet of Things (IoT), and its connections to real-life applications at work. It provides a detailed guidance on how FinTech

digitalizes business operations, improves productivity and efficiency, and optimizes resource management with the help of some new concepts, such as AIOps, MLOps and DevSecOps. Readers will also discover how FinTech Innovations connect BFSI to the rest of the world with growing interests in Open Banking, Banking-as-a-Service (BaaS) and FinTech-as-a-Service (FaaS). To help readers understand

how FinTech has unlocked numerous opportunities for tapping into the massive substantial group of customers, this book illustrates the massive changes already underway and provides insights into changes yet to come through practical examples and applications with illustrative figures and summary tables, making this book a handy quick reference for all things of

<p>FinTech.Related Link(s) <i>Field Guide to Hadoop</i> Morgan & Claypool Publishers This is the first book to offer a comprehensive yet concise overview of the challenges and opportunities presented by the use of big data in healthcare. The respective chapters address a range of aspects: from health management to patient safety; from the human factor perspective to ethical and</p>	<p>economic considerations, and many more. By providing a historical background on the use of big data, and critically analyzing current approaches together with issues and challenges related to their applications, the book not only sheds light on the problems entailed by big data, but also paves the way for possible solutions and future research directions.</p>	<p>Accordingly, it offers an insightful reference guide for health information technology professionals, healthcare managers, healthcare practitioners, and patients alike, aiding them in their decision-making processes; and for students and researchers whose work involves data science-related research issues in healthcare. <u>Professional Hadoop Solutions</u></p>
---	--	--

Harvard Business Review Press The volume LNCS 12393 constitutes the papers of the 22nd International Conference Big Data Analytics and Knowledge Discovery which will be held online in September 2020. The 15 full papers presented together with 14 short papers plus 1 position paper in this volume were carefully reviewed and selected from a total of 77 submissions. This volume offers a wide

range to following subjects on theoretical and practical aspects of big data analytics and knowledge discovery as a new generation of big data repository, data pre-processing, data mining, text mining, sequences, graph mining, and parallel processing. Handbook of Big Data Technologies "O'Reilly Media, Inc." Hadoop in Action teaches readers how to use Hadoop and write

MapReduce programs. The intended readers are programmers, architects, and project managers who have to process large amounts of data offline. Hadoop in Action will lead the reader from obtaining a copy of Hadoop to setting it up in a cluster and writing data analytic programs. The book begins by making the basic idea of Hadoop and MapReduce easier to grasp by applying the

default Hadoop installation to a few easy-to-follow tasks, such as analyzing changes in word frequency across a body of documents. The book continues through the basic concepts of MapReduce applications developed using Hadoop, including a close look at framework components, use of Hadoop for a variety of data analysis tasks, and numerous examples of Hadoop in action.

Hadoop in Action will explain how to use Hadoop and present design patterns and practices of programming MapReduce. MapReduce is a complex idea both conceptually and in its implementation, and Hadoop users are challenged to learn all the knobs and levers for running Hadoop. This book takes you beyond the mechanics of running Hadoop, teaching you to write meaningful

programs in a MapReduce framework. This book assumes the reader will have a basic familiarity with Java, as most code examples will be written in Java. Familiarity with basic statistical concepts (e.g. histogram, correlation) will help the reader appreciate the more advanced data processing examples. Purchase of the print book comes with an offer of a free PDF, ePub,

and Kindle eBook from Manning. Also available is all code from the book.

Data-intensive Text Processing with MapReduce

"O'Reilly Media, Inc." This book is a compendium of the proceedings of the International Conference on Big-Data and Cloud Computing. The papers discuss the recent advances in the areas of big data analytics, data analytics in

cloud, smart cities and grid, etc. This volume primarily focuses on the application of knowledge which promotes ideas for solving problems of the society through cutting-edge big-data technologies. The essays featured in this proceeding provide novel ideas that contribute for the growth of world class research and development. It will be useful to researchers in

the area of advanced engineering sciences.

Intelligence in Big Data Technologies –Beyond the Hype

"O'Reilly Media, Inc." This book constitutes the refereed proceedings of the 14th International Conference on Software Reuse for Dynamic Systems in the Cloud and Beyond, ICSR 2015, held in Miami, FL, USA, in January 2015. The 21 revised full papers presented together with

3 revised short papers were carefully reviewed and selected from 60 submissions. The papers cover several software engineering areas where software reuse is important, such as software product lines, domain analysis, open source, components, cloud, quality.

Big Data, Big Challenges: A Healthcare Perspective

Springer
Nature
Summary
Mahout in Action is a

hands-on introduction to machine learning with Apache Mahout.

Following real-world examples, the book presents practical use cases and then illustrates how Mahout can be applied to solve them.

Includes a free audio- and video-enhanced ebook. About the

Technology A computer system that learns and adapts as it collects data can be really powerful. Mahout,

Apache's open source machine learning project, captures the core algorithms of recommendation systems, classification, and clustering in ready-to-use, scalable libraries. With Mahout, you can immediately apply to your own projects the machine learning techniques that drive Amazon, Netflix, and others. About this Book This book covers machine learning using Apache

Mahout. Based on experience with real-world applications, it introduces practical use cases and illustrates how Mahout can be applied to solve them. It places particular focus on issues of scalability and how to apply these techniques against large data sets using the Apache Hadoop framework. This book is written for developers familiar with Java -- no prior experience with Mahout is assumed. Owners of a Manning pBook purchased anywhere in the world can download a free eBook from manning.com at any time. They can do so multiple times and in any or all formats available (PDF, ePub or Kindle). To do so, customers must register their printed copy on Manning's site by creating a user account and then following instructions printed on the pBook registration insert at the front of the book. What's Inside Use group data to make individual recommendations Find logical clusters within your data Filter and refine with on-the-fly classification Free audio and video extras Table of Contents Meet Apache Mahout PART 1 RECOMMENDATIONS Introducing recommenders Representing

recommender data Making recommendations Taking recommender suggestions to production Distributing recommendations on computations	to classification Training a classifier Evaluating and tuning a classifier Deploying a classifier Case study: Shop It To Me	International Conference on Computing, Power and Communication Technologies 2019 (GUCON 2019), organized by Galgotias University, India, in September 2019. The content is divided into three sections – data mining and big data analysis, communication technologies, and cloud computing and computer networks. In-depth discussions of various issues within these
PART 2 CLUSTERING Introduction to clustering Representing data Clustering algorithms in Mahout Evaluating and improving clustering quality Taking clustering to production Real-world applications of clustering	<u>Hadoop: The Definitive Guide</u> KHANNA PUBLISHING Apache Mahout <i>Artificial Intelligence: Concepts, Methodologies, Tools, and Applications</i> Springer Nature This book gathers selected high-quality papers presented at the	
PART 3 CLASSIFICATION Introduction		

broad areas provide an intriguing and insightful reference guide for researchers, engineers and students alike. Big Data Using Hadoop and Hive Springer Nature Ready to unlock the power of your data? With this comprehensive guide, you'll learn how to build and maintain reliable, scalable, distributed systems with Apache Hadoop. This book is ideal for programmers

looking to analyze datasets of any size, and for administrators who want to set up and run Hadoop clusters. You'll find illuminating case studies that demonstrate how Hadoop is used to solve specific problems. This third edition covers recent changes to Hadoop, including material on the new MapReduce API, as well as MapReduce 2 and its more flexible execution

model (YARN). Store large datasets with the Hadoop Distributed File System (HDFS) Run distributed computations with MapReduce Use Hadoop's data and I/O building blocks for compression, data integrity, serialization (including Avro), and persistence Discover common pitfalls and advanced features for writing real-world MapReduce programs Design, build, and

administer a dedicated Hadoop cluster—or run Hadoop in the cloud Load data from relational databases into HDFS, using Sqoop Perform large-scale data processing with the Pig query language Analyze datasets with Hive, Hadoop’s data warehousing system Take advantage of HBase for structured and semi-structured data, and ZooKeeper for building distributed

systems
[Apache Mahout Essentials](#)
 Simon and Schuster
 Our world is being revolutionized by data-driven methods: access to large amounts of data has generated new insights and opened exciting new opportunities in commerce, science, and computing applications. Processing the enormous quantities of data necessary for these advances requires large clusters,

making distributed computing paradigms more crucial than ever. MapReduce is a programming model for expressing distributed computations on massive datasets and an execution framework for large-scale data processing on clusters of commodity servers. The programming model provides an easy-to-understand abstraction for designing scalable algorithms,

while the execution framework transparently handles many system-level details, ranging from scheduling to synchronization to fault tolerance. This book focuses on MapReduce algorithm design, with an emphasis on text processing algorithms common in natural language processing, information retrieval, and machine learning. We introduce the notion of MapReduce design

patterns, which represent general reusable solutions to commonly occurring problems across a variety of problem domains. This book not only intends to help the reader "think in MapReduce", but also discusses limitations of the programming model as well. This volume is a printed version of a work that appears in the Synthesis Digital Library

of Engineering and Computer Science. Synthesis Lectures provide concise, original presentations of important research and development topics, published quickly, in digital and print formats. For more information visit www.morganclaypool.com *Apache Mahout* Academic Press Summary Hadoop in Practice, Second Edition provides over

100 tested, instantly useful techniques that will help you conquer big data, using Hadoop. This revised new edition covers changes and new features in the Hadoop core architecture, including MapReduce 2. Brand new chapters cover YARN and integrating Kafka, Impala, and Spark SQL with Hadoop. You'll also get new and updated techniques for Flume, Sqoop, and Mahout,

all of which have seen major new versions recently. In short, this is the most practical, up-to-date coverage of Hadoop available anywhere. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Book It's always a good time to upgrade your Hadoop skills! Hadoop in Practice, Second Edition provides a

collection of 104 tested, instantly useful techniques for analyzing real-time streams, moving data securely, machine learning, managing large-scale clusters, and taming big data using Hadoop. This completely revised edition covers changes and new features in Hadoop core, including MapReduce 2 and YARN. You'll pick up hands-on best practices for integrating Spark, Kafka, and Impala

<p>with Hadoop, and get new and updated techniques for the latest versions of Flume, Sqoop, and Mahout. In short, this is the most practical, up-to-date coverage of Hadoop available. Readers need to know a programming language like Java and have basic familiarity with Hadoop. What's Inside Thoroughly updated for Hadoop 2 How to write YARN applications Integrate real-time technologies</p>	<p>like Storm, Impala, and Spark Predictive analytics using Mahout and RR Readers need to know a programming language like Java and have basic familiarity with Hadoop. About the Author Alex Holmes works on tough big-data problems. He is a software engineer, author, speaker, and blogger specializing in large-scale Hadoop projects. Table of Contents PART 1</p>	<p>BACKGROUND AND FUNDAMENTALS Hadoop in a heartbeat Introduction to YARN PART 2 DATA LOGISTICS Data serialization—working with text and beyond Organizing and optimizing data in HDFS Moving data into and out of Hadoop PART 3 BIG DATA PATTERNS Applying MapReduce patterns to big data Utilizing data structures and algorithms at scale Tuning, debugging, and testing</p>
--	--	---

PART 4
BEYOND
MAPREDUCE
SQL on
Hadoop
Writing a
YARN
application
*Database
Systems for
Advanced
Applications*
CRC Press
Engineering
education
methods and
standards are
important
features of
engineering
programs that
should be
carefully
designed both
to provide
students and
stakeholders
with valuable,
active,
integrated
learning
experiences,

and to provide
a vehicle for
assessing
program
outcomes.
With the
driving force
of the
globalization
of the
engineering
profession,
standards
should be
developed for
mutual
recognition of
engineering
education
across the
world, but it is
proving
difficult to
achieve. The
Handbook of
Research on
Engineering
Education in a
Global Context
provides
innovative
insights into

the
importance of
quality
training and
preparation
for
engineering
students. It
explores the
common and
current
problems
encountered
in areas such
as quality and
standards,
management
information
systems,
innovation
and enhanced
learning
technologies
in education,
as well as the
challenges of
employability,
entrepreneurs
hip, and
diversity. This
publication is
vital reference

source for science and engineering educators, engineering professionals, and educational administrators interested in topics centered on the education of students in the field of engineering.

Hadoop 2 Quick-Start Guide O'Reilly Media Apache Mahout is a scalable machine learning library with algorithms for clustering, classification, and recommendations. It

empowers users to analyze patterns in large, diverse, and complex datasets faster and more scalably. This book is an all-inclusive guide to analyzing large and complex datasets using Apache Mahout. It explains complicated but very effective machine learning algorithms simply, in relation to real-world practical examples. Starting from

the fundamental concepts of machine learning and Apache Mahout, this book guides you through Apache Mahout's implementations of machine learning techniques including classification, clustering, and recommendations. During this exciting walkthrough, real-world applications, a diverse range of popular algorithms and their implementations, code examples,

evaluation strategies, and best practices are given for each technique. Finally, you will learn vdata visualization techniques for Apache Mahout to bring your data to life.

Machine Learning and the Internet of Medical Things in Healthcare

Springer
Machine Learning and the Internet of Medical Things in Healthcare discusses the applications and challenges of

machine learning for healthcare applications. The book provides a platform for presenting machine learning-enabled healthcare techniques and offers a mathematical and conceptual background of the latest technology. It describes machine learning techniques along with the emerging platform of the Internet of Medical Things used by practitioners

and researchers worldwide. The book includes deep feed forward networks, regularization, optimization algorithms, convolutional networks, sequence modeling, and practical methodology. It also presents the concepts of the Internet of Things, the set of technologies that develops traditional devices into smart devices. Finally, the book offers research perspectives, covering the

convergence of machine learning and IoT. It also presents the application of these technologies in the development of healthcare frameworks. Provides an introduction to the Internet of Medical Things through the principles and applications of machine learning. Explains the functions and applications of machine learning in various applications such as ultrasound imaging,

biomedical signal processing, robotics, and biomechanics. Includes coverage of the evolution of healthcare applications with machine learning, including Clinical Decision Support Systems, artificial intelligence in biomedical engineering, and AI-enabled connected health informatics, supported by real-world case studies. **The Web at Graduation and Beyond**

Springer
The go-to guidebook for deploying Big Data solutions with Hadoop. Today's enterprise architects need to understand how the Hadoop frameworks and APIs fit together, and how they can be integrated to deliver real-world solutions. This book is a practical, detailed guide to building and implementing those solutions, with code-level instruction in the

popular Wrox tradition. It covers storing data with HDFS and Hbase, processing data with MapReduce, and automating data processing with Oozie. Hadoop security, running Hadoop with Amazon Web Services, best practices, and automating Hadoop processes in real time are also covered in depth. With in-depth code examples in Java and XML and the latest on recent additions to

the Hadoop ecosystem, this complete resource also covers the use of APIs, exposing their inner workings and allowing architects and developers to better leverage and customize them. The ultimate guide for developers, designers, and architects who need to build and deploy Hadoop applications. Covers storing and processing data with various technologies, automating data

processing, Hadoop security, and delivering real-time solutions. Includes detailed, real-world examples and code-level guidelines. Explains when, why, and how to use these tools effectively. Written by a team of Hadoop experts in the programmer-to-programmer style. Professional Hadoop Solutions is the reference enterprise architects and developers

need to maximize the power of Hadoop.

Software Reuse for Dynamic Systems in the Cloud and Beyond

Springer

The growth of a global digital economy has enabled rapid communication, instantaneous movement of funds, and availability of vast amounts of information. With this come challenges such as the vulnerability of digitalized sociotechnological systems (STSs) to

destructive events (earthquakes, disease events, terrorist attacks). Similar issues arise for disruptions to complex linked natural and social systems (from changing climates, evolving urban environments, etc.). This book explores new approaches to the resilience of sociotechnological and natural-social systems in a digital world of big data, extraordinary

computing capacity, and rapidly developing methods of Artificial Intelligence. Most of the book's papers were presented at the Workshop on Big Data and Systems Analysis held at the International Institute for Applied Systems Analysis in Laxenburg, Austria in February, 2020. Their authors are associated with the Task Group "Advanced mathematical tools for data-

driven applied systems analysis” created and sponsored by CODATA in November, 2018. The world-wide COVID-19 pandemic illustrates the vulnerability of our healthcare systems, supply chains, and social infrastructure, and confronts our notions of what makes a system resilient. We have found that use of AI tools can lead to problems when unexpected events occur. On the other

hand, the vast amounts of data available from sensors, satellite images, social media, etc. can also be used to make modern systems more resilient. Papers in the book explore disruptions of complex networks and algorithms that minimize departure from a previous state after a disruption; introduce a multigrammatical framework for the technological and resource bases of today’s large-

scale industrial systems and the transformations resulting from disruptive events; and explain how robotics can enhance pre-emptive measures or post-disaster responses to increase resiliency. Other papers explore current directions in data processing and handling and principles of FAIRness in data; how the availability of large amounts of data can aid in the

development of resilient STSs and challenges to overcome in doing so. The book also addresses interactions between humans and built environments, focusing on how AI can inform today's smart and connected buildings and make them resilient, and how AI tools can increase resilience to misinformation and its dissemination. *Big Data Analytics Beyond Hadoop* Springer

There is an easier way to build Hadoop applications. With this hands-on book, you'll learn how to use Cascading, the open source abstraction framework for Hadoop that lets you easily create and manage powerful enterprise-grade data processing applications—without having to learn the intricacies of MapReduce. Working with sample apps based on Java and other JVM

languages, you'll quickly learn Cascading's streamlined approach to data processing, data filtering, and workflow optimization. This book demonstrates how this framework can help your business extract meaningful information from large amounts of distributed data. Start working on Cascading example projects right away Model and analyze unstructured data in any

format, from any source Build and test applications with familiar constructs and reusable components Work with the Scalding and Cascalog Domain-Specific Languages Easily deploy applications to Hadoop, regardless of cluster location or data size Build workflows that integrate several big data frameworks and processes Explore common use cases for Cascading, including

features and tools that support them Examine a case study that uses a dataset from the Open Data Initiative [Future And Fintech, The: Abcdi And Beyond](#) Springer If your team is investigating ways to design applications for the cloud, this concise book introduces 11 architecture patterns that can help you take advantage of cloud-platform services. You'll learn how each of

these platform-agnostic patterns work, when they might be useful in the cloud, and what impact they'll have on your application architecture. You'll also see an example of each pattern applied to an application built with Windows Azure. The patterns are organized into four major topics, such as scalability and handling failure, and primer chapters provide background

on each topic. With the information in this book, you'll be able to make informed decisions for designing effective cloud-native applications that maximize the value of cloud services, while also paying attention to user experience and operational efficiency. Learn about architectural patterns for: Scalability. Discover the advantages of horizontal scaling. Patterns

covered include Horizontally Scaling Compute, Queue-Centric Workflow, and Auto-Scaling. Big data. Learn how to handle large amounts of data across a distributed system. Eventual consistency is explained, along with the MapReduce and Database Sharding patterns. Handling failure. Understand how multitenant cloud services and commodity hardware

influence your applications. Patterns covered include Busy Signal and Node Failure. Distributed users. Learn how to overcome delays due to network latency when building applications for a geographically distributed user base. Patterns covered include Colocation, Valet Key, CDN, and Multi-Site Deployment. **Intelligent Systems** Apache MahoutApach

<p>e Mahout: Beyond MapReduce. Distributed algorithm design This book is about designing mathematical and Machine Learning algorithms using the Apache Mahout "Samsara" platform. The material takes on best programming practices as well as conceptual approaches to attacking Machine Learning problems in big datasets. Math is explained, followed by</p>	<p>code examples of distributed and in- memory computations. Written by Apache Mahout committers for people who want to learn how to design distributed math algorithms as well as how to use some of the new Mahout "Samsara" algorithms off- the-shelf. The book covers Apache Mahout 0.10 and 0.11. Handboo k of E-Tourism This book constitutes the</p>	<p>proceedings of the International Conference on E-business and Strategy, iCETS 2012, held in Tianjin, China, in August 2012. The 65 revised full papers presented were carefully reviewed and selected from 231 submissions. The papers feature contemporary research on developments in the fields of e-business technology, information management systems, and business strategy. Topics</p>
--	---	---

addressed are	commerce	e-business
latest	applications,	education,
development	social	entrepreneurs
on e-business	networking	hip and e-
technology,	and social	learning,
computer	engineering	digital
science and	for e-business,	economy
software	e-business	strategy, as
engineering	strategic	well as
for e-business,	management	internet and
e-business	and	e-commerce
and e-	economics	policy.
	development,	

Best Sellers - Books :

- [Fast Like A Girl: A Woman's Guide To Using The Healing Power Of Fasting To Burn Fat, Boost Energy, And Balance Hormones By Dr. Mindy Pelz](#)
- [Atomic Habits: An Easy & Proven Way To Build Good Habits & Break Bad Ones](#)
- [I Love You Like No Otter: A Funny And Sweet Board Book For Babies And Toddlers \(punderland\) By Rose Rossner](#)
- [We'll Always Have Summer \(the Summer I Turned Pretty\) By Jenny Han](#)
- [Dog Man: Twenty Thousand Fleas Under The Sea: A Graphic Novel \(dog Man #11\): From The Creator Of Captain Underpants By Dav Pilkey](#)
- [Young Forever: The Secrets To Living Your Longest, Healthiest Life \(the Dr. Hyman Library, 11\) By Dr. Mark Hyman Md](#)
- [Rich Dad Poor Dad: What The Rich Teach Their](#)

Kids About Money That The Poor And Middle Class Do Not!

- The Creative Act: A Way Of Being By Rick Rubin
- A Court Of Thorns And Roses (a Court Of Thorns And Roses, 1) By Sarah J. Maas
- The Last Thing He Told Me: A Novel