Storing And Managing Big Data Nosql Hadoop And More High Impact Strategies What You Need To Know Definitions Adoptions Impact Benefits Matur Kevin Roebuck

In today's fast growing digital world, the web, mobile, social networks and other digital platforms are producing enormous amounts of data that hold intelligence and valuable information. Correctly used it has the power to create sustainable value in different forms for businesses. The commonly used term for this data is Big Data, which includes structured, unstructured and hybrid structured data. However, Big Data is of limited value unless insightful information can be extracted from the sources of data. The solution is Big Data analytics, and how managers and executives can capture value from this vast resource of information and insights. This book develops a simple framework and a non-technical approach to help the reader understand, digest and analyze data, and produce meaningful analytics to make informed decisions. It will support value creation within businesses, from customer care to product innovation, from sales and marketing to operational performance. The authors provide multiple case studies on global industries and business units, chapter summaries and discussion questions for the reader to consider and explore. Big Data for Managers also presents small cases and challenges for the reader to work on – making this a thorough and practical guide for students and managers.

This open access book is part of the LAMBDA Project (Learning, Applying, Multiplying Big Data Analytics), funded by the European Union, GA No. 809965. Data Analytics involves applying algorithmic processes to derive insights. Nowadays it is used in many industries to allow organizations and companies to make better decisions as well as to verify or disprove existing theories or models. The term data analytics is often used interchangeably with intelligence, statistics, reasoning, data mining, knowledge discovery, and others. The goal of this book is to introduce some of the definitions, methods, tools, frameworks, and solutions for big data processing, starting from the process of information extraction and knowledge representation, via knowledge processing and analytics to visualization, sense-making, and practical applications. Each chapter in this book addresses some pertinent aspect of the data processing chain, with a specific focus on understanding Enterprise Knowledge Graphs, Semantic Big Data Architectures, and Smart Data Analytics solutions. This book is addressed to graduate students from technical disciplines, to professional audiences following continuous education short courses, and to researchers from diverse areas following self-study courses. Basic skills in computer science, mathematics, and statistics are required.

The data lake is a daring new approach for harnessing the power of big data technology and providing convenient self-service capabilities. But is it right for your company? This book is based on discussions with practitioners and executives from more than a hundred organizations, ranging from data-driven companies such as Google, LinkedIn, and Facebook, to governments and traditional corporate enterprises. You'll learn what a data lake is, why enterprises need one, and how to build one successfully with the best practices in this book. Alex Gorelik, CTO and founder of Waterline Data, explains why old systems and processes can no longer support data needs in the enterprise. Then, in a collection of essays about data lake implementation, you'll examine data lake initiatives, analytic projects, experiences, and best practices from data experts working in various industries. Get a succinct introduction to data warehousing, big data, and data science Learn various paths enterprises take to build a data lake Explore how to build a self-service model and best practices for providing analysts access to the

data Use different methods for architecting your data lake Discover ways to implement a data lake from experts in different industries

This report improves the evidence base on the role of Data Driven Innovation for promoting growth and well-being, and provide policy guidance on how to maximise the benefits of DDI and mitigate the associated economic and societal risks.

Strategy is Digital

Computing Systems and Approaches

Big Data Business Guide

Issues and Challenges

Big Data Technologies and Applications

Managing Data in Motion

From the Foreword: "Big Data Management and Processing is [a] state-of-the-art book that deals with a wide range of topical themes in the field of Big Data. The book, which probes many issues related to this exciting and rapidly growing field, covers processing, management, analytics, and applications... [It] is a very valuable addition to the literature. It will serve as a source of up-to-date research in this continuously developing area. The book also provides an opportunity for researchers to explore the use of advanced computing technologies and their impact on enhancing our capabilities to conduct more sophisticated studies." ---Sartaj Sahni, University of Florida, USA "Big Data Management and Processing covers the latest Big Data research results in processing, analytics, management and applications. Both fundamental insights and representative applications are provided. This book is a timely and valuable resource for students, researchers and seasoned practitioners in Big Data fields. --Hai Jin, Huazhong University of Science and Technology, China Big Data Management and Processing explores a range of big data related issues and their impact on the design of new computing systems. The twenty-one chapters were carefully selected and feature contributions from several outstanding researchers. The book endeavors to strike a balance between theoretical and practical coverage of innovative problem solving techniques for a range of platforms. It serves as a repository of paradigms, technologies, and applications that target different facets of big data computing systems. The first part of the book explores energy and resource management issues, as well as legal compliance and quality management for Big Data. It covers In-Memory computing and In-Memory data grids, as well as co-scheduling for high performance computing applications. The second part of the book includes comprehensive coverage of Hadoop and Spark, along with security, privacy, and trust challenges and solutions. The latter part of the book covers mining and clustering in Big Data, and includes applications in genomics, hospital big data processing, and vehicular cloud computing. The book also analyzes funding for Big Data projects. This IBM® RedpaperTM provides a reference architecture, based on Apache Hadoop, to help businesses gain control over their data, meet tight service level agreements (SLAs) around their data applications, and turn data-driven insight into effective action. Big Data Networked Storage Solution for Hadoop delivers

the capabilities for ingesting, storing, and managing large data sets with high reliability. IBM InfoSphere® Big InsightsTM provides an innovative analytics platform that processes and analyzes all types of data to turn large complex data into insight. IBM InfoSphere BigInsights brings the power of Hadoop to the enterprise. With built-in analytics, extensive integration capabilities, and the reliability, security and support that you require, IBM can help put your big data to work for you. This IBM Redpaper publication provides basic guidelines and best practices for how to size and configure Big Data Networked Storage Solution for Hadoop.

Healthcare transformation requires us to continually look at new and better ways to manage insights - both within and outside the organization today. Increasingly, the ability to glean and operationalize new insights efficiently as a byproduct of an organization's day-to-day operations is becoming vital to hospitals and health systems ability to survive and prosper. One of the long-standing challenges in healthcare informatics has been the ability to deal with the sheer variety and volume of disparate healthcare data and the increasing need to derive veracity and value out of it. Demystifying Big Data and Machine Learning for Healthcare investigates how healthcare organizations can leverage this tapestry of big data to discover new business value, use cases, and knowledge as well as how big data can be woven into pre-existing business intelligence and analytics efforts. This book focuses on teaching you how to: Develop skills needed to identify and demolish big-data myths Become an expert in separating hype from reality Understand the V's that matter in healthcare and why Harmonize the 4 C's across little and big data Choose data fi delity over data quality Learn how to apply the NRF Framework Master applied machine learning for healthcare Conduct a guided tour of learning algorithms Recognize and be prepared for the future of artificial intelligence in healthcare via best practices, feedback loops, and contextually intelligent agents (CIAs) The variety of data in healthcare spans multiple business workflows, formats (structured, un-, and semi-structured), integration at point of care/need, and integration with existing knowledge. In order to deal with these realities, the authors propose new approaches to creating a knowledge-driven learning organization-based on new and existing strategies, methods and technologies. This book will address the long-standing challenges in healthcare informatics and provide pragmatic recommendations on how to deal with them.

Data analytics is core to business and decision making. The rapid increase in data volume, velocity and variety offers both opportunities and challenges. While open source solutions to store big data, like Hadoop, offer platforms for exploring value and insight from big data, they were not originally developed with data security and governance in mind. Big Data Management discusses numerous policies, strategies and recipes for managing big data. It addresses data security, privacy, controls and life cycle management offering modern principles and open source architectures for successful governance of big data. The author has collected best practices from the world's leading organizations that have

successfully implemented big data platforms. The topics discussed cover the entire data management life cycle, data quality, data stewardship, regulatory considerations, data council, architectural and operational models are presented for successful management of big data. The book is a must-read for data scientists, data engineers and corporate leaders who are implementing big data platforms in their organizations.

Techniques and Technologies in Geoinformatics

Storage, Sharing, and Security

Managing Big Data in Cloud Computing Environments

Principles of Database Management

Processing and Management

Big Data Management

The best-selling author of Big Data is back, this time with a unique and in-depth insight into how specific companies use big data. Big data is on the tip of everyone's tongue. Everyone understands its power and importance, but many fail to grasp the actionable steps and resources required to utilise it effectively. This book fills the knowledge gap by showing how major companies are using big data every day, from an up-close, on-the-ground perspective. From technology, media and retail, to sport teams, government agencies and financial institutions, learn the actual strategies and processes being used to learn about customers, improve manufacturing, spur innovation, improve safety and so much more. Organised for easy dip-in navigation, each chapter follows the same structure to give you the information you need quickly. For each company profiled, learn what data was used, what problem it solved and the processes put it place to make it practical, as well as the technical details, challenges and lessons learned from each unique scenario. Learn how predictive analytics helps Amazon, Target, John Deere and Apple understand their customers Discover how big data is behind the success of Walmart, LinkedIn, Microsoft and more Learn how big data is changing medicine, law enforcement, hospitality, fashion, science and banking Develop your own big data strategy by accessing additional reading materials at the end of each chapter

Large Scale and Big Data: Processing and Management provides readers with a central source of reference on the data management techniques currently available for large-scale data processing. Presenting chapters written by leading researchers, academics, and practitioners, it addresses the fundamental challenges associated with Big Data processing tools and techniques across a range of computing environments. The book begins by discussing the basic concepts and tools of

large-scale Big Data processing and cloud computing. It also provides an overview of different programming models and cloud-based deployment models. The book's second section examines the usage of advanced Big Data processing techniques in different domains, including semantic web, graph processing, and stream processing. The third section discusses advanced topics of Big Data processing such as consistency management, privacy, and security. Supplying a comprehensive summary from both the research and applied perspectives, the book covers recent research discoveries and applications, making it an ideal reference for a wide range of audiences, including researchers and academics working on databases, data mining, and web scale data processing. After reading this book, you will gain a fundamental understanding of how to use Big Data-processing tools and techniques effectively across application domains. Coverage includes cloud data management architectures, big data analytics visualization, data management, analytics for vast amounts of unstructured data, clustering, classification, link analysis of big data, scalable data mining, and machine learning techniques.

Managing Big Data is a simple book which introduces students and professionals to Big Data. Although the book has been designed for unassisted reading, lot of insights from the author makes this a very thoughtful book which will automatically lead to yearning for more learning on the subject.

Although there are already some books published on Big Data, most of them only cover basic concepts and society impacts and ignore the internal implementation details-making them unsuitable to R&D people. To fill such a need, Big Data: Storage, Sharing, and Security examines Big Data management from an R&D perspective. It covers the 3S desi

Delivering the Promise of Big Data and Data Science

Large Scale and Big Data

Big Data

The Next Frontier for Innovation, Competition, and Productivity

Storing, Managing, and Protecting Digital Information in Classic, Virtualized, and Cloud Environments

Big Data Security

Big data has always been a major challenge in geoinformatics as geospatial data come in various types and formats, new geospatial data are acquired very fast, and geospatial databases are inherently very large. And while there have been advances in hardware and

software for handling big data, they often fall short of handling geospatial big data efficiently and effectively. Big Data: Techniques and Technologies in Geoinformatics tackles these challenges head on, integrating coverage of techniques and technologies for storing, managing, and computing geospatial big data. Providing a perspective based on analysis of time, applications, and resources, this book familiarizes readers with geospatial applications that fall under the category of big data. It explores new trends in geospatial data collection, such as geo-crowdsourcing and advanced data collection technologies such as LiDAR point clouds. The book features a range of topics on big data techniques and technologies in geoinformatics including distributed computing, geospatial data analytics, social media, and volunteered geographic information. With chapters contributed by experts in geoinformatics and in domains such as computing and engineering, the book provides an understanding of the challenges and issues of big data in geoinformatics applications. The book is a single collection of current and emerging techniques, technologies, and tools that are needed to collect, analyze, manage, process, and visualize geospatial big data.

Data mining of massive data sets is transforming the way we think about crisis response, marketing, entertainment, cybersecurity and national intelligence. Collections of documents, images, videos, and networks are being thought of not merely as bit strings to be stored, indexed, and retrieved, but as potential sources of discovery and knowledge, requiring sophisticated analysis techniques that go far beyond classical indexing and keyword counting, aiming to find relational and semantic interpretations of the phenomena underlying the data. Frontiers in Massive Data Analysis examines the frontier of analyzing massive amounts of data, whether in a static database or streaming through a system. Data at that scale--terabytes and petabytes--is increasingly common in science (e.g., particle physics, remote sensing, genomics), Internet commerce, business analytics, national security, communications, and elsewhere. The tools that work to infer knowledge from data at smaller scales do not necessarily work, or work well, at such massive scale. New tools, skills, and approaches are necessary, and this report identifies many of them, plus promising research directions to explore. Frontiers in

Massive Data Analysis discusses pitfalls in trying to infer knowledge from massive data, and it characterizes seven major classes of computation that are common in the analysis of massive data. Overall, this report illustrates the cross-disciplinary knowledge--from computer science, statistics, machine learning, and application disciplines--that must be brought to bear to make useful inferences from massive data.

Big data is currently one of the most critical emerging technologies. Organizations around the world are looking to exploit the explosive growth of data to unlock previously hidden insights in the hope of creating new revenue streams, gaining operational efficiencies, and obtaining greater understanding of customer needs. It is important to think of big data and analytics together. Big data is the term used to describe the recent explosion of different types of data from disparate sources. Analytics is about examining data to derive interesting and relevant trends and patterns, which can be used to inform decisions, optimize processes, and even drive new business models. With today's deluge of data comes the problems of processing that data, obtaining the correct skills to manage and analyze that data, and establishing rules to govern the data's use and distribution. The big data technology stack is ever growing and sometimes confusing, even more so when we add the complexities of setting up big data environments with large up-front investments. Cloud computing seems to be a perfect vehicle for hosting big data workloads. However, working on big data in the cloud brings its own challenge of reconciling two contradictory design principles. Cloud computing is based on the concepts of consolidation and resource pooling, but big data systems (such as Hadoop) are built on the shared nothing principle, where each node is independent and self-sufficient. A solution architecture that can allow these mutually exclusive principles to coexist is required to truly exploit the elasticity and ease-of-use of cloud computing for big data environments. This IBM® RedpaperTM publication is aimed at chief architects, line-of-business executives, and CIOs to provide an understanding of the cloud-related challenges they face and give prescriptive guidance for how to realize the benefits of big data solutions quickly and cost-effectively.

This book is a wonderful collection of chapters that posits how managers need to cope in

the Big Data era. It highlights many of the emerging developments in technologies, applications, and trends related to management's needs in this Big Data era. —Dr. Jay Liebowitz, Harrisburg University of Science and Technology This book presents some meaningful work on Big Data analytics and its applications. Each chapter generates helpful guidance to the readers on Big Data analytics and its applications, challenges, and prospects that is necessary for organizational strategic direction. —Dr. Alex Koohang, Middle Georgia State University Big Data is a concept that has caught the attention of practitioners, academicians, and researchers. Big Data offers organizations the possibility of gaining a competitive advantage by managing, collecting, and analyzing massive amounts of data. As the promises and challenges posed by Big Data have increased over the past decade, significant issues have developed regarding how data can be used for improving management. Big Data can be understood as large amounts of data generated by the Internet and a variety of connected smart devices and sensors. This book discusses the main challenges posed by Big Data in a manner relevant to both practitioners and scholars. It examines how companies can leverage Big Data analytics to act and optimize the business. This book brings together the theory and practice of management in the era of Big Data. It offers a look at the current state of Big Data, including a comprehensive overview of both research and practical applications. By bringing together conceptual thinking and empirical research on the nature, meaning, and development of Big Data in management, this book unifies research on Big Data in management to stimulate new directions for academic investigation as well as practice.

Concepts, Technology, and Architecture

Knowledge Graphs and Big Data Processing

New Horizons for a Data-Driven Economy

Managing Big Data Integration in the Public Sector

Big Data For Dummies

A Roadmap for Usage and Exploitation of Big Data in Europe

Disk-Based Algorithms for Big Data is a product of recent advances in the areas of big data, data analytics, and the underlying file systems and data management algorithms used to support the storage and analysis of massive data collections. The book

discusses hard disks and their impact on data management, since Hard Disk Drives continue to be common in large data clusters. It also explores ways to store and retrieve data though primary and secondary indices. This includes a review of different in-memory sorting and searching algorithms that build a foundation for more sophisticated on-disk approaches like mergesort, B-trees, and extendible hashing. Following this introduction, the book transitions to more recent topics, including advanced storage technologies like solid-state drives and holographic storage; peer-to-peer (P2P) communication; large file systems and query languages like Hadoop/HDFS, Hive, Cassandra, and Presto; and NoSQL databases like Neo4j for graph structures and MongoDB for unstructured document data. Designed for senior undergraduate and graduate students, as well as professionals, this book is useful for anyone interested in understanding the foundations and advances in big data storage and management, and big data analytics. About the Author Dr. Christopher G. Healey is a tenured Professor in the Department of Computer Science and the Goodnight Distinguished Professor of Analytics in the Institute for Advanced Analytics, both at North Carolina State University in Raleigh, North Carolina. He has published over 50 articles in major journals and conferences in the areas of visualization, visual and data analytics, computer graphics, and artificial intelligence. He is a recipient of the National Science Foundation’s CAREER Early Faculty Development Award and the North Carolina State University Outstanding Instructor Award. He is a Senior Member of the Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE), and an Associate Editor of ACM Transaction on Applied Perception, the leading worldwide journal on the application of human perception to issues in computer science.

Cloud computing has proven to be a successful paradigm of service-oriented computing, and has revolutionized the way computing infrastructures are abstracted and used. By means of cloud computing technology, massive data can be managed effectively and efficiently to support various aspects of problem solving and decision making. Managing Big Data in Cloud Computing Environments explores the latest advancements in the area of data management and analysis in the cloud. Providing timely, research-based information relating to data storage, sharing, extraction, and indexing in cloud systems, this publication is an ideal reference source for graduate students, IT specialists, researchers, and professionals working in the areas of data and knowledge engineering.

Managing Data in Motion describes techniques that have been developed for significantly reducing the complexity of managing system interfaces and enabling scalable architectures. Author April Reeve brings over two decades of experience to present a vendor-neutral approach to moving data between computing environments and systems. Readers will learn the techniques, technologies, and best practices for managing the passage of data between computer systems and integrating disparate data together in an enterprise environment. The average enterprise's computing environment is comprised of hundreds to thousands computer systems that have been built, purchased, and acquired over time. The data from these various systems needs to be integrated for reporting and analysis, shared for business transaction processing, and converted from one format to another when old systems are replaced and new systems are acquired. The management of the "data in motion" in

organizations is rapidly becoming one of the biggest concerns for business and IT management. Data warehousing and conversion, real-time data integration, and cloud and "big data" applications are just a few of the challenges facing organizations and businesses today. Managing Data in Motion tackles these and other topics in a style easily understood by business and IT managers as well as programmers and architects. Presents a vendor-neutral overview of the different technologies and techniques for moving data between computer systems including the emerging solutions for unstructured as well as structured data types Explains, in non-technical terms, the architecture and components required to perform data integration Describes how to reduce the complexity of managing system interfaces and enable a scalable data architecture that can handle the dimensions of "Big Data"

The digital age has presented an exponential growth in the amount of data available to individuals looking to draw conclusions based on given or collected information across industries. Challenges associated with the analysis, security, sharing, storage, and visualization of large and complex data sets continue to plague data scientists and analysts alike as traditional data processing applications struggle to adequately manage big data. The Handbook of Research on Big Data Storage and Visualization Techniques is a critical scholarly resource that explores big data analytics and technologies and their role in developing a broad understanding of issues pertaining to the use of big data in multidisciplinary fields. Featuring coverage on a broad range of topics, such as architecture patterns, programing systems, and computational energy, this publication is geared towards professionals, researchers, and students seeking current research and application topics on the subject.

Big Data Computing

Information Storage and Management

Data Governance Principles for Big Data Analytics

Handbook of Research on Big Data Storage and Visualization Techniques

A Beginner's Guide

How Companies Can Use Big Data in the Value Chain

The book features research papers presented at the International Conference on Emerging Technologies in Data Mining and Information Security (IEMIS 2018) held at the University of Engineering & Management, Kolkata, India, on February 23-25, 2018. It comprises high-quality research by academics and industrial experts in the field of computing and communication, including full-length papers, research-in-progress papers, case studies related to all the areas of data mining, machine learning, IoT and information security.

Find the right big data solution for your business ororganization Big data management is one of the major challenges facingbusiness, industry, and not-for-profit organizations. Data setssuch as customer transactions for a mega-retailer, weather patternsmonitored by meteorologists, or social network activity can quicklyoutpace the capacity of traditional data management tools. If youneed to develop or manage big data solutions, you'll appreciate howthese four experts define, explain, and guide you through this newand often confusing concept. You'll learn what it is, why itmatters, and how to choose and implement solutions that work. Effectively managing big data is an issue of growing importanceto businesses, not-for-profit organizations, government, and ITprofessionals Authors are experts in information

management, big data, and avariety of solutions Explains big data in detail and discusses how to select andimplement a solution, security concerns to consider, data storageand presentation issues, analytics, and much more Provides essential information in a no-nonsense,easy-to-understand style that is empowering Big Data For Dummies cuts through the confusion and helpsyou take charge of big data solutions for your organization.

The era of rapidly progressing technology we live in generates vast amounts of data; however, the challenge exists in understanding how to aggressively monitor and make sense of this data. Without a better understanding of how to collect and manage such large data sets, it becomes increasingly difficult to successfully utilize them. Managing Big Data Integration in the Public Sector is a pivotal reference source for the latest scholarly research on the application of big data analytics in government contexts and identifies various strategies in which big data platforms can generate improvements within that sector. Highlighting issues surrounding data management, current models, and real-world applications, this book is ideally designed for professionals, government agencies, researchers, and non-profit organizations interested in the benefits of big data analytics applied in the public sphere.

Traditional intrusion detection and logfile analysis are no longer enough to protect today ' s complex networks. In this practical guide, security researcher Michael Collins shows you several techniques and tools for collecting and analyzing network traffic datasets. You ' ll understand how your network is used, and what actions are necessary to protect and improve it. Divided into three sections, this book examines the process of collecting and organizing data, various tools for analysis, and several different analytic scenarios and techniques. It ' s ideal for network administrators and operational security analysts familiar with scripting. Explore network, host, and service sensors for capturing security data Store data traffic with relational databases, graph databases, Redis, and Hadoop Use SiLK, the R language, and other tools for analysis and visualization Detect unusual phenomena through Exploratory Data Analysis (EDA) Identify significant structures in networks with graph analysis Determine the traffic that ' s crossing service ports in a network Examine traffic volume and behavior to spot DDoS and database raids Get a step-by-step process for network mapping and inventory

Management in the Era of Big Data

Big Data Networked Storage Solution for Hadoop

Big Data Management and Processing

Network Security Through Data Analysis

Disk-Based Algorithms for Big Data

Building Big Data and Analytics Solutions in the Cloud

Process Safety and Big Data discusses the principles of process safety and advanced information technologies. It explains how these principles are applied to the process industry and provides examples of applications in process safety control and decision support systems. This book helps to address problems that researchers face in industry that are the result of increased process complexity and that have an impact on safety issues. It shows ways to tackle these safety issues by implementing modern information technologies, such as big data analysis and artificial intelligence. It provides an integrated approach to modern information technologies used in control and management of process safety in industry. The book also considers indicators and criteria in effective safety decisions, and addresses the issue of how

big data would provide support for improved, autonomous, data-driven decisions. Paves the way for the digital transformation of safety science and safety management Takes a system approach to advanced information technologies used in process safety Applies big data technologies to process safety Includes multiple pertinent case studies

Data Warehousing in the Age of the Big Data will help you and your organization make the most of unstructured data with your existing data warehouse. As Big Data continues to revolutionize how we use data, it doesn't have to create more confusion. Expert author Krish Krishnan helps you make sense of how Big Data fits into the world of data warehousing in clear and concise detail. The book is presented in three distinct parts. Part 1 discusses Big Data, its technologies and use cases from early adopters. Part 2 addresses data warehousing, its shortcomings, and new architecture options, workloads, and integration techniques for Big Data and the data warehouse. Part 3 deals with data governance, data visualization, information life-cycle management, data scientists, and implementing a Big Data–ready data warehouse. Extensive appendixes include case studies from vendor implementations and a special segment on how we can build a healthcare information factory. Ultimately, this book will help you navigate through the complex layers of Big Data and data warehousing while providing you information on how to effectively think about using all these technologies and the architectures to design the next-generation data warehouse. Learn how to leverage Big Data by effectively integrating it into your data warehouse. Includes real-world examples and use cases that clearly demonstrate Hadoop, NoSQL, HBASE, Hive, and other Big Data technologies Understand how to optimize and tune your current data warehouse infrastructure and integrate newer infrastructure matching data processing workloads and requirements Introductory, theory-practice balanced text teaching the fundamentals of databases to advanced undergraduates or graduate students in information systems or computer science.

Big data analytics and cloud computing is the fastest growing technologies in current era. This text book serves as a purpose in providing an understanding of big data principles and framework at the beginner?s level. The text book covers various essential concepts of big-data analytics and processing tools such as HADOOP and YARN. The Textbook covers an analogical understanding on bridging cloud computing with big-data technologies with essential cloud infrastructure protocol and ecosystem concepts. PART I: Hadoop Distributed File System Basics, Running Example Programs and Benchmarks, Hadoop MapReduce Framework Essential Hadoop Tools, Hadoop YARN Applications, Managing Hadoop with Apache Ambari, Basic Hadoop Administration Procedures PART II: Introduction to Cloud Computing: Origins and Influences, Basic Concepts and Terminology, Goals and Benefits, Risks and Challenges. Fundamental Concepts and Models: Roles and Boundaries, Cloud Characteristics, Cloud Delivery Models, Cloud Deployment Models. Cloud Computing Technologies:Broadband networks and internet architecture, data center technology, virtualization technology, web technology, multi-tenant technology, service Technology Cloud Infrastructure

Mechanisms:Logical Network Perimeter, Virtual Server, Cloud Storage Device, Cloud Usage Monitor, Resource Replication, Ready-made environment

Building Situational Awareness

Data-Driven Innovation Big Data for Growth and Well-Being

Frontiers in Massive Data Analysis

Managing Big Data

Big Data Analytics and Cloud Computing

How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results

The objective of this book is to introduce the basic concepts of big data computing and then to describe the total solution of big data problems using HPCC, an open-source computing platform. The book comprises 15 chapters broken into three parts. The first part, Big Data Technologies, includes introductions to big data concepts and techniques; big data analytics; and visualization and learning techniques. The second part, LexisNexis Risk Solution to Big Data, focuses on specific technologies and techniques developed at LexisNexis to solve critical problems that use big data analytics. It covers the open source High Performance Computing Cluster (HPCC Systems®) platform and its architecture, as well as parallel data languages ECL and KEL, developed to effectively solve big data problems. The third part, Big Data Applications, describes various data intensive applications solved on HPCC Systems. It includes applications such as cyber security, social network analytics including fraud, Ebola spread modeling using big data analytics, unsupervised learning, and image classification. The book is intended for a wide variety of people including researchers, scientists, programmers, engineers, designers, developers, educators, and students. This book can also be beneficial for business managers, entrepreneurs, and investors.

THE SERIES: FRONTIERS IN COMPUTATIONAL INTELLIGENCE The series Frontiers In Computational Intelligence is envisioned to provide comprehensive coverage and understanding of cutting edge research in computational intelligence. It intends to augment the scholarly discourse on all topics relating to the advances in artifi cial life and machine learning in the form of metaheuristics, approximate reasoning, and robotics. Latest research

findings are coupled with applications to varied domains of engineering and computer sciences. This field is steadily growing especially with the advent of novel machine learning algorithms being applied to different domains of engineering and technology. The series brings together leading researchers that intend to continue to advance the field and create a broad knowledge about the most recent research. Series Editor Dr. Siddhartha Bhattacharyya, CHRIST (Deemed to be University), Bangalore, India Editorial Advisory Board Dr. Elizabeth Behrman, Wichita State University, Kansas, USA Dr. Goran Klepac Dr. Leo Mrsic, Algebra University College, Croatia Dr. Aboul Ella Hassanien, Cairo University, Egypt Dr. Jan Platos, VSB-Technical University of Ostrava, Czech Republic Dr. Xiao-Zhi Gao, University of Eastern Finland, Finland Dr. Wellington Pinheiro dos Santos, Federal University of Pernambuco, Brazil

In this book readers will find technological discussions on the existing and emerging technologies across the different stages of the big data value chain. They will learn about legal aspects of big data, the social impact, and about education needs and requirements. And they will discover the business perspective and how big data technology can be exploited to deliver value within different sectors of the economy. The book is structured in four parts: Part I "The Big Data Opportunity" explores the value potential of big data with a particular focus on the European context. It also describes the legal, business and social dimensions that need to be addressed, and briefly introduces the European Commission's BIG project. Part II "The Big Data Value Chain" details the complete big data lifecycle from a technical point of view, ranging from data acquisition, analysis, curation and storage, to data usage and exploitation. Next, Part III "Usage and Exploitation of Big Data" illustrates the value creation possibilities of big data applications in various sectors, including industry, healthcare, finance, energy, media and public services. Finally, Part IV "A Roadmap for Big Data Research" identifies and prioritizes the cross-sectorial requirements for big data research, and outlines the most urgent and challenging technological, economic, political and societal issues for big data in Europe. This compendium summarizes more than two years of work performed by a leading group of major European research centers and industries in the

context of the BIG project. It brings together research findings, forecasts and estimates related to this challenging technological context that is becoming the major axis of the new digitally transformed business environment.

Big data has presented a number of opportunities across industries. With these opportunities come a number of challenges associated with handling, analyzing, and storing large data sets. One solution to this challenge is cloud computing, which supports a massive storage and computation facility in order to accommodate big data processing. Managing and Processing Big Data in Cloud Computing explores the challenges of supporting big data processing and cloud-based platforms as a proposed solution. Emphasizing a number of crucial topics such as data analytics, wireless networks, mobile clouds, and machine learning, this publication meets the research needs of data analysts, IT professionals, researchers, graduate students, and educators in the areas of data science, computer programming, and IT development.

Creating Value

The Practical Guide to Storing, Managing and Analyzing Big and Small Data

Big Data in Practice

Process Safety and Big Data

Big Data for Managers

Managing and Processing Big Data in Cloud Computing

This book presents strategies and practices to allow everyday companies to cope with the fundamentally changing landscape of business models and to take advantage of the huge business opportunities arising from the advent of big data. It develops several case studies from companies in traditional industries like LEGO, Yamato and Mediq, but also examines small start-ups like Space Tango, which is partnering with major multinationals to develop new business models using big data. The book argues that businesses need to adapt and embark on their big data journey, helps them take the first step, and guides them along their way. It presents successful examples and deducts essential takeaway lessons from them, equipping executives to capitalize on big data and enabling them to make intelligent decisions in the big data transformation, giving their companies an essential competitive edge.

This book presents a detailed review of high-performance computing infrastructures for next-generation big data and fast data analytics. Features: includes case studies and learning activities throughout the book and self-study exercises in every chapter; presents detailed case studies on social media analytics for intelligent businesses and on big data analytics (BDA) in the healthcare sector; describes the network

infrastructure requirements for effective transfer of big data, and the storage infrastructure requirements of applications which generate big data; examines real-time analytics solutions; introduces in-database processing and in-memory analytics techniques for data mining; discusses the use of mainframes for handling real-time big data and the latest types of data management systems for BDA; provides information on the use of cluster, grid and cloud computing systems for BDA; reviews the peer-to-peer techniques and tools and the common information visualization techniques, used in BDA.

Learn Big Data from the ground up with this complete and up-to-date resource from leaders in the field Big Data: Concepts, Technology, and Architecture delivers a comprehensive treatment of Big Data tools, terminology, and technology perfectly suited to a wide range of business professionals, academic researchers, and students. Beginning with a fulsome overview of what we mean when we say, “Big Data,” the book moves on to discuss every stage of the lifecycle of Big Data. You’ll learn about the creation of structured, unstructured, and semi-structured data, data storage solutions, traditional database solutions like SQL, data processing, data analytics, machine learning, and data mining. You’ll also discover how specific technologies like Apache Hadoop, SQOOP, and Flume work. Big Data also covers the central topic of big data visualization with Tableau, and you’ll learn how to create scatter plots, histograms, bar, line, and pie charts with that software. Accessibly organized, Big Data includes illuminating case studies throughout the material, showing you how the included concepts have been applied in real-world settings. Some of those concepts include: The common challenges facing big data technology and technologists, like data heterogeneity and incompleteness, data volume and velocity, storage limitations, and privacy concerns Relational and non-relational databases, like RDBMS, NoSQL, and NewSQL databases Virtualizing Big Data through encapsulation, partitioning, and isolating, as well as big data server virtualization Apache software, including Hadoop, Cassandra, Avro, Pig, Mahout, Oozie, and Hive The Big Data analytics lifecycle, including business case evaluation, data preparation, extraction, transformation, analysis, and visualization Perfect for data scientists, data engineers, and database managers, Big Data also belongs on the bookshelves of business intelligence analysts who are required to make decisions based on large volumes of information. Executives and managers who lead teams responsible for keeping or understanding large datasets will also benefit from this book.

Due to market forces and technological evolution, Big Data computing is developing at an increasing rate. A wide variety of novel approaches and tools have emerged to tackle the challenges of Big Data, creating both more opportunities and more challenges for students and professionals in the field of data computation and analysis. Presenting a mix of industry cases and theory, Big Data Computing discusses the technical and practical issues related to Big Data in intelligent information management. Emphasizing the adoption and diffusion of Big Data tools and technologies in industry, the book introduces a broad range of Big Data concepts, tools, and techniques. It covers a wide range of research, and provides comparisons between state-of-the-art approaches. Comprised of five sections, the book focuses on: What Big Data is and why it is important Semantic technologies Tools and methods Business and economic perspectives Big Data applications across industries

Data Warehousing in the Age of Big Data

Emerging Technologies in Data Mining and Information Security

High-Performance Big-Data Analytics

Demystifying Big Data and Machine Learning for Healthcare
Big Data for Growth and Well-Being
Proceedings of IEMIS 2018, Volume 2

The new edition of a bestseller, now revised and update throughout! This new edition of the unparalleled bestseller serves as a full training course all in one and as the world's largest data storage company, EMC is the ideal author for such a critical resource. They cover the components of a storage system and the different storage system models while also offering essential new material that explores the advances in existing technologies and the emergence of the "Cloud" as well as updates and vital information on new technologies. Features a separate section on emerging area of cloud computing Covers new technologies such as: data de-duplication, unified storage, continuous data protection technology, virtual provisioning, FCoE, flash drives, storage tiering, big data, and more Details storage models such as Network Attached Storage (NAS), Storage Area Network (SAN), Object Based Storage along with virtualization at various infrastructure components Explores Business Continuity and Security in physical and virtualized environment Includes an enhanced Appendix for additional information This authoritative guide is essential for getting up to speed on the newest advances in information storage and management.

The Enterprise Big Data Lake
Data Integration Best Practice Techniques and Technologies