

highlighting key concepts and definitions, important checkpoints, communication activities, and best practices. * A companion Web site includes links to numerous data quality resources, including many of the planning and information-gathering templates featured in the text, quick summaries of key ideas from the Ten Step methodology, and other tools and information available online.

A thorough reference on database administration outlines a variety of DBA roles and responsibilities and discusses such topics as data modeling and normalization, database/application design, change management, database security and data integrity, performance issues, disaster planning, and other essentials. Original. (Advanced)

A key task that any aspiring data-driven organization needs to learn is data wrangling, the process of converting raw data into something truly useful. This practical guide provides business analysts with an overview of various data wrangling techniques and tools, and puts the practice of data wrangling into context by asking, "What are you trying to do and why?" Wrangling data consumes roughly 50-80% of an analyst's time before any kind of analysis is possible. Written by key executives at Trifacta, this book walks you through the wrangling process by exploring several factors—time, granularity, scope, and structure—that you need to consider as you begin to work with data. You'll learn a shared language and a comprehensive understanding of data wrangling, with an emphasis on recent agile analytic processes used by many of today's data-driven organizations. Appreciate the importance—and the satisfaction—of wrangling data the right way. Understand what kind of data is available Choose which data to use and at what level of detail Meaningfully combine multiple sources of data Decide how to distill the results to a size and shape that can drive downstream analysis

How to implement social technology in business, spur collaborative innovation and drive winning programs to improve products, services, and long-term profits and growth. The road to social media marketing is now well paved: A July 2009 Anderson Analytics study found 60% of the Internet population uses social networks and social media sites such as Facebook, MySpace, and Twitter. Collaboration and innovation, driven by social technology, are "what's next." Written by the author of the bestselling Social Media Marketing: An Hour a Day in collaboration with Jake McKee, Social Media Marketing: The Next Generation of Business Engagement takes marketers, product managers, small business owners, senior executives and organizational leaders on to the next step in social technology and its application in business. In particular, this book explains how to successfully implement a variety tools, how to ensure higher levels of customer engagement, and how to build on the lessons learned and information gleaned from first-generation social media marketing efforts and to carry this across your organization. This book: Details how to develop, implement, monitor and measure successful social media activities, and how to successfully act on feedback from the social web Discusses conversation-monitoring tools and platforms to accelerate the business innovation cycle along with the metrics required to prove the success of social technology adoption Connects the social dots more deeply across the entire organization, moving beyond marketing and into product development, customer service and customer-driven innovation, and the benefits of encouraging employee collaboration. Social media has become a central component of marketing: Collaborative, social technology is now moving across the organization, into business functions ranging from HR and legal to product management and the supply chain. Social Media Marketing: The Next Generation of Business Engagement is the perfect book for marketers, business unit managers and owners, HR professionals and anyone else looking to better understand how to use social technologies and platforms to build loyalty in customers, employees, partners and suppliers to drive long term growth and profits.

Database Design and Implementation

Learn Informatica in 24 Hours

The Complete Project Lifecycle For Decision-support Applications

Internals and Design Principles

Ten Steps to Quality Data and Trusted Information (TM)

DAMA-DMBOK

Practical Techniques for Data Preparation

This book examines the nature of retail financial transaction infrastructures. Contributions assume a long-term outlook in their exploration of the key financial processes and systems that support a global transition to a cashless economy. The volume offers both modern and historic accounts that demonstrate the constantly changing role of payment instruments. It brings together different theoretical approaches to the study, re-examining and forecasting changes in retail payment systems. Chapters explore a global transition to a cashless society and contemplate future alternatives to cash, cheques and plastic, featuring the perspectives of academics from different disciplines in conversation and industry participants from six continents. Readers are invited to discover the innovation in payment systems and how it co-evolves with changes in society and organisations through personal, corporate and governmental processes.

Next-Generation Big DataA Practical Guide to Apache Kudu, Impala, and SparkApress

This is the eBook of the printed book and may not include any media, website access codes, or print supplements that may come packaged with the bound book. Database Systems: The Complete Book is ideal for Database Systems and Database Design and Application courses offered at the junior, senior and graduate levels in Computer Science departments. A basic understanding of algebraic expressions and laws, logic, basic data structure, OOP concepts, and programming environments is implied. Written by well-known computer scientists, this introduction to database systems offers a comprehensive approach, focusing on database design, database use, and implementation of database applications and database management systems. The first half of the book provides in-depth coverage of databases from the point of view of the database designer, user, and application programmer. It covers the latest database standards SQL:1999, SQL/PSM, SQL/CLI, JDBC, ODL, and XML, with broader coverage of SQL than most other texts. The second half of the book provides in-depth coverage of databases from the point of view of the DBMS implementor. It focuses on storage structures, query processing, and transaction management. The book covers the main techniques in these areas with broader coverage of query optimization than most other texts, along with advanced topics including multidimensional and bitmap indexes, distributed transactions, and information integration techniques.

Invented about 40 years ago and called ubiquitous less than 10 years later. B-tree indexes have been used in a wide variety of computing systems from handheld devices to mainframes and server farms. Over the years, many techniques have been added to the basic design in order to improve efficiency or to add functionality. Examples include separation of updates to structure or contents, utility operations such as non-logged yet transactional index creation, and robust query processing such as graceful degradation during index-to-index navigation. Modern B-Tree Techniques reviews the basics of B-trees and of B-tree indexes in databases, transactional techniques and query processing techniques related to B-trees, B-tree utilities essential for database operations, and many optimizations and improvements. It is intended both as a tutorial and as a reference, enabling researchers to compare index innovations with advanced B-tree techniques and enabling professionals to select features, functions, and tradeoffs most appropriate for their data management challenges.

Big Data

The Antivirus Hacker's Handbook

Getting Started with Talend Open Studio for Data Integration

Biggs

Hadoop: The Definitive Guide

Modern B-Tree Techniques

Find the right big data solution for your business or organization Big data management is one of the major challenges facingbusiness, industry, and not-for-profit organizations. Data sets such 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.

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 quality of a data warehouse (DWH) is the elusive aspect of it, not because it is hard to achieve (once we agree what it is), but because it is difficult to describe. We propose the notion that quality is not an attribute or a feature that a product has to possess, but rather a relationship between that product and each and every stakeholder. More specifically, the relationship between the software quality and the organization that produces the products is explored. Quality of data that populates the DWH is the main concern of the book, therefore we propose a definition for data quality as: "fitness to serve each and every purpose". Methods are proposed throughout the book to help readers achieve data warehouse quality.

This book constitutes the thoroughly refereed proceedings of the 14th International Conference on Metadata and Semantic Research, MTSR 2020, held in Madrid, Spain, in December 2020. Due to the COVID-19 pandemic the conference was held online. The 24 full and 13 short papers presented were carefully reviewed and selected from 82 submissions. The papers are organized in the following tracks: metadata, linked data, semantics and ontologies; metadata and semantics for digital libraries, information retrieval, big, linked, social and open data; metadata and semantics for agriculture, food, and environment, AgroSEN 2020; metadata and semantics for open repositories, research information systems and data infrastructures; digital humanities and digital curation, DHC 2020; metadata and semantics for cultural collections and applications; european and national projects; knowledge IT artifacts (KITA) in professional communities and aggregations, KITA 2020.

Google BigQuery: The Definitive Guide

The Definitive Guide