

Executing Data Quality Projects Ten Steps To Quality Data And Trusted Information

This is likewise one of the factors by obtaining the soft documents of this **Executing Data Quality Projects Ten Steps To Quality Data And Trusted Information** by online. You might not require more era to spend to go to the books foundation as skillfully as search for them. In some cases, you likewise get not discover the proclamation Executing Data Quality Projects Ten Steps To Quality Data And Trusted Information that you are looking for. It will certainly squander the time.

However below, past you visit this web page, it will be consequently certainly simple to get as competently as download lead Executing Data Quality Projects Ten Steps To Quality Data And Trusted Information

It will not understand many mature as we tell before. You can accomplish it even if feat something else at house and even in your workplace. in view of that easy! So, are you question? Just exercise just what we have the funds for below as without difficulty as evaluation **Executing Data Quality Projects Ten Steps To Quality Data And Trusted Information** what you subsequently to read!

Data Governance - John Ladley 2019-11-08

Managing data continues to grow as a necessity for modern organizations. There are seemingly infinite opportunities for organic growth, reduction of costs, and creation of new products and services. It has become apparent that none of these opportunities can happen smoothly without data governance. The cost of exponential data growth and privacy / security concerns are becoming burdensome. Organizations will encounter unexpected consequences in new sources of risk. The solution to these challenges is also data governance; ensuring balance between risk and opportunity. Data Governance, Second Edition, is for any executive, manager or data professional who needs to understand or implement a data governance program. It is required to ensure consistent, accurate and reliable data across their organization. This book offers an overview of why data governance is needed, how to design, initiate, and execute a

program and how to keep the program sustainable. This valuable resource provides comprehensive guidance to beginning professionals, managers or analysts looking to improve their processes, and advanced students in Data Management and related courses. With the provided framework and case studies all professionals in the data governance field will gain key insights into launching successful and money-saving data governance program. Incorporates industry changes, lessons learned and new approaches Explores various ways in which data analysts and managers can ensure consistent, accurate and reliable data across their organizations Includes new case studies which detail real-world situations Explores all of the capabilities an organization must adopt to become data driven Provides guidance on various approaches to data governance, to determine whether an organization should be low profile, central controlled, agile, or traditional Provides guidance on using

technology and separating vendor hype from sincere delivery of necessary capabilities Offers readers insights into how their organizations can improve the value of their data, through data quality, data strategy and data literacy Provides up to 75% brand-new content compared to the first edition

Data Quality Assessment - Arkady Maydanchik
2007-04-01

Imagine a group of prehistoric hunters armed with stone-tipped spears. Their primitive weapons made hunting large animals, such as mammoths, dangerous work. Over time, however, a new breed of hunters developed. They would stretch the skin of a previously killed mammoth on the wall and throw their spears, while observing which spear, thrown from which angle and distance, penetrated the skin the best. The data gathered helped them make better spears and develop better hunting strategies. Quality data is the key to any advancement, whether it's from the Stone Age to the Bronze

Age. Or from the Information Age to whatever Age comes next. The success of corporations and government institutions largely depends on the efficiency with which they can collect, organize, and utilize data about products, customers, competitors, and employees. Fortunately, improving your data quality doesn't have to be such a mammoth task. DATA QUALITY ASSESSMENT is a must read for anyone who needs to understand, correct, or prevent data quality issues in their organization. Skipping theory and focusing purely on what is practical and what works, this text contains a proven approach to identifying, warehousing, and analyzing data errors - the first step in any data quality program. Master techniques in:

- Data profiling and gathering metadata
- Identifying, designing, and implementing data quality rules
- Organizing rule and error catalogues
- Ensuring accuracy and completeness of the data quality assessment
- Constructing the dimensional data quality scorecard
- Executing a recurrent data

quality assessment This is one of those books that marks a milestone in the evolution of a discipline. Arkady's insights and techniques fuel the transition of data quality management from art to science -- from crafting to engineering. From deep experience, with thoughtful structure, and with engaging style Arkady brings the discipline of data quality to practitioners. David Wells, Director of Education, Data Warehousing Institute

Driven by Data - Paul Bambrick-Santoyo
2010-04-12

Offers a practical guide for improving schools dramatically that will enable all students from all backgrounds to achieve at high levels. Includes assessment forms, an index, and a DVD.

Leading Change - John P. Kotter 2012

Offers advice on how to lead an organization into change, including establishing a sense of urgency, developing a vision and strategy, and generating short-term wins.

Occupational Outlook Handbook - United

States. Bureau of Labor Statistics 1976

Data Cleaning - Ihab F. Ilyas 2019-06-18

Data quality is one of the most important problems in data management, since dirty data often leads to inaccurate data analytics results and incorrect business decisions. Poor data across businesses and the U.S. government are reported to cost trillions of dollars a year. Multiple surveys show that dirty data is the most common barrier faced by data scientists. Not surprisingly, developing effective and efficient data cleaning solutions is challenging and is rife with deep theoretical and engineering problems. This book is about data cleaning, which is used to refer to all kinds of tasks and activities to detect and repair errors in the data. Rather than focus on a particular data cleaning task, we give an overview of the end-to-end data cleaning process, describing various error detection and repair methods, and attempt to anchor these proposals with multiple taxonomies and views.

Specifically, we cover four of the most common and important data cleaning tasks, namely, outlier detection, data transformation, error repair (including imputing missing values), and data deduplication. Furthermore, due to the increasing popularity and applicability of machine learning techniques, we include a chapter that specifically explores how machine learning techniques are used for data cleaning, and how data cleaning is used to improve machine learning models. This book is intended to serve as a useful reference for researchers and practitioners who are interested in the area of data quality and data cleaning. It can also be used as a textbook for a graduate course. Although we aim at covering state-of-the-art algorithms and techniques, we recognize that data cleaning is still an active field of research and therefore provide future directions of research whenever appropriate.

The Practitioner's Guide to Data Quality Improvement - David Loshin 2010-11-22

The Practitioner's Guide to Data Quality Improvement offers a comprehensive look at data quality for business and IT, encompassing people, process, and technology. It shares the fundamentals for understanding the impacts of poor data quality, and guides practitioners and managers alike in socializing, gaining sponsorship for, planning, and establishing a data quality program. It demonstrates how to institute and run a data quality program, from first thoughts and justifications to maintenance and ongoing metrics. It includes an in-depth look at the use of data quality tools, including business case templates, and tools for analysis, reporting, and strategic planning. This book is recommended for data management practitioners, including database analysts, information analysts, data administrators, data architects, enterprise architects, data warehouse engineers, and systems analysts, and their managers. Offers a comprehensive look at data quality for business and IT, encompassing

people, process, and technology. Shows how to institute and run a data quality program, from first thoughts and justifications to maintenance and ongoing metrics. Includes an in-depth look at the use of data quality tools, including business case templates, and tools for analysis, reporting, and strategic planning.

Data Quality - Thomas C. Redman 2001

Can any subject inspire less excitement than "data quality"? Yet a moment's thought reveals the ever-growing importance of quality data. From restated corporate earnings, to incorrect prices on the web, to the bombing of the Chinese Embassy, the media reports the impact of poor data quality on a daily basis. Every business operation creates or consumes huge quantities of data. If the data are wrong, time, money, and reputation are lost. In today's environment, every leader, every decision maker, every operational manager, every consumer, indeed everyone has a vested interest in data quality. *Data Quality: The Field Guide* provides the

practical guidance needed to start and advance a data quality program. It motivates interest in data quality, describes the most important data quality problems facing the typical organization, and outlines what an organization must do to improve. It consists of 36 short chapters in an easy-to-use field guide format. Each chapter describes a single issue and how to address it. The book begins with sections that describe why leaders, whether CIOs, CFOs, or CEOs, should be concerned with data quality. It explains the pros and cons of approaches for addressing the issue. It explains what those organizations with the best data do. And it lays bare the social issues that prevent organizations from making headway. "Field tips" at the end of each chapter summarize the most important points. Allows readers to go directly to the topic of interest Provides web-based material so readers can cut and paste figures and tables into documents within their organizations Gives step-by-step instructions for applying most techniques and

summarizes what "works"

The Anarchist Cookbook - William Powell

2018-03-11

The Anarchist Cookbook will shock, it will disturb, it will provoke. It places in historical perspective an era when "Turn on, Burn down, Blow up" are revolutionary slogans of the day. Says the author "This book... is not written for the members of fringe political groups, such as the Weatherman, or The Minutemen. Those radical groups don't need this book. They already know everything that's in here. If the real people of America, the silent majority, are going to survive, they must educate themselves. That is the purpose of this book." In what the author considers a survival guide, there is explicit information on the uses and effects of drugs, ranging from pot to heroin to peanuts. There is detailed advice concerning electronics, sabotage, and surveillance, with data on everything from bugs to scramblers. There is a comprehensive chapter on natural, non-lethal,

and lethal weapons, running the gamut from cattle prods to sub-machine guns to bows and arrows.

Measuring Data Quality for Ongoing Improvement - Laura Sebastian-Coleman

2012-12-31

The Data Quality Assessment Framework shows you how to measure and monitor data quality, ensuring quality over time. You'll start with general concepts of measurement and work your way through a detailed framework of more than three dozen measurement types related to five objective dimensions of quality: completeness, timeliness, consistency, validity, and integrity. Ongoing measurement, rather than one time activities will help your organization reach a new level of data quality. This plain-language approach to measuring data can be understood by both business and IT and provides practical guidance on how to apply the DQAF within any organization enabling you to prioritize measurements and effectively report on results.

Strategies for using data measurement to govern and improve the quality of data and guidelines for applying the framework within a data asset are included. You'll come away able to prioritize which measurement types to implement, knowing where to place them in a data flow and how frequently to measure. Common conceptual models for defining and storing of data quality results for purposes of trend analysis are also included as well as generic business requirements for ongoing measuring and monitoring including calculations and comparisons that make the measurements meaningful and help understand trends and detect anomalies. Demonstrates how to leverage a technology independent data quality measurement framework for your specific business priorities and data quality challenges Enables discussions between business and IT with a non-technical vocabulary for data quality measurement Describes how to measure data quality on an ongoing basis with generic

measurement types that can be applied to any situation

Ten Steps to a Results-Based Monitoring and Evaluation System - Jody Zall Kusek 2004-06-15
This Handbook provides a comprehensive ten-step model that will help guide development practitioners through the process of designing and building a results-based monitoring and evaluation system.

Data Stewardship - David Plotkin 2013-09-16
Data stewards in business and IT are the backbone of a successful data governance implementation because they do the work to make a company's data trusted, dependable, and high quality. Data Stewardship explains everything you need to know to successfully implement the stewardship portion of data governance, including how to organize, train, and work with data stewards, get high-quality business definitions and other metadata, and perform the day-to-day tasks using a minimum of the steward's time and effort. David Plotkin has

loaded this book with practical advice on stewardship so you can get right to work, have early successes, and measure and communicate those successes, gaining more support for this critical effort. Provides clear and concise practical advice on implementing and running data stewardship, including guidelines on how to organize based on company structure, business functions, and data ownership Shows how to gain support for your stewardship effort, maintain that support over the long-term, and measure the success of the data stewardship effort and report back to management Includes detailed lists of responsibilities for each type of data steward and strategies to help the Data Governance Program Office work effectively with the data stewards

Data Driven - Thomas C. Redman 2008-09-22

Your company's data has the potential to add enormous value to every facet of the organization -- from marketing and new product development to strategy to financial

management. Yet if your company is like most, it's not using its data to create strategic advantage. Data sits around unused -- or incorrect data fouls up operations and decision making. In *Data Driven*, Thomas Redman, the "Data Doc," shows how to leverage and deploy data to sharpen your company's competitive edge and enhance its profitability. The author reveals:

- The special properties that make data such a powerful asset
- The hidden costs of flawed, outdated, or otherwise poor-quality data
- How to improve data quality for competitive advantage
- Strategies for exploiting your data to make better business decisions
- The many ways to bring data to market
- Ideas for dealing with political struggles over data and concerns about privacy rights

Your company's data is a key business asset, and you need to manage it aggressively and professionally. Whether you're a top executive, an aspiring leader, or a product-line manager, this eye-opening book provides the tools and thinking you need to do that.

R for Data Science - Hadley Wickham

2016-12-12

Learn how to use R to turn raw data into insight, knowledge, and understanding. This book introduces you to R, RStudio, and the tidyverse, a collection of R packages designed to work together to make data science fast, fluent, and fun. Suitable for readers with no previous programming experience, R for Data Science is designed to get you doing data science as quickly as possible. Authors Hadley Wickham and Garrett Golemund guide you through the steps of importing, wrangling, exploring, and modeling your data and communicating the results. You'll get a complete, big-picture understanding of the data science cycle, along with basic tools you need to manage the details. Each section of the book is paired with exercises to help you practice what you've learned along the way. You'll learn how to:

Wrangle—transform your datasets into a form convenient for analysis
Program—learn powerful

R tools for solving data problems with greater clarity and ease
Explore—examine your data, generate hypotheses, and quickly test them
Model—provide a low-dimensional summary that captures true "signals" in your dataset
Communicate—learn R Markdown for integrating prose, code, and results

The 4 Disciplines of Execution - Chris

McChesney 2016-04-12

BUSINESS STRATEGY. "The 4 Disciplines of Execution" offers the what but also how effective execution is achieved. They share numerous examples of companies that have done just that, not once, but over and over again. This is a book that every leader should read! (Clayton Christensen, Professor, Harvard Business School, and author of "The Innovator's Dilemma)." Do you remember the last major initiative you watched die in your organization? Did it go down with a loud crash? Or was it slowly and quietly suffocated by other competing priorities? By the time it finally disappeared, it s

likely no one even noticed. What happened? The whirlwind of urgent activity required to keep things running day-to-day devoured all the time and energy you needed to invest in executing your strategy for tomorrow. "The 4 Disciplines of Execution" can change all that forever.

Executing Data Quality Projects - Danette McGilvray 2008

Infonomics - Douglas B. Laney 2017-09-05

Many senior executives talk about information as one of their most important assets, but few behave as if it is. They report to the board on the health of their workforce, their financials, their customers, and their partnerships, but rarely the health of their information assets. Corporations typically exhibit greater discipline in tracking and accounting for their office furniture than their data. *Infonomics* is the theory, study, and discipline of asserting economic significance to information. It strives to apply both economic and asset management principles and practices

to the valuation, handling, and deployment of information assets. This book specifically shows: CEOs and business leaders how to more fully wield information as a corporate asset CIOs how to improve the flow and accessibility of information CFOs how to help their organizations measure the actual and latent value in their information assets. More directly, this book is for the burgeoning force of chief data officers (CDOs) and other information and analytics leaders in their valiant struggle to help their organizations become more info-savvy. Author Douglas Laney has spent years researching and developing *Infonomics* and advising organizations on the infinite opportunities to monetize, manage, and measure information. This book delivers a set of new ideas, frameworks, evidence, and even approaches adapted from other disciplines on how to administer, wield, and understand the value of information. *Infonomics* can help organizations not only to better develop, sell,

and market their offerings, but to transform their organizations altogether. "Doug Laney masterfully weaves together a collection of great examples with a solid framework to guide readers on how to gain competitive advantage through what he labels "the unruly asset" - data. The framework is comprehensive, the advice practical and the success stories global and across industries and applications." Liz Rowe, Chief Data Officer, State of New Jersey "A must read for anybody who wants to survive in a data centric world." Shaun Adams, Head of Data Science, Betterbathrooms.com "Phenomenal! An absolute must read for data practitioners, business leaders and technology strategists. Doug's lucid style has a set a new standard in providing intelligible material in the field of information economics. His passion and knowledge on the subject exudes thru his literature and inspires individuals like me." Ruchi Rajasekhar, Principal Data Architect, MISO Energy "I highly recommend Infonomics

to all aspiring analytics leaders. Doug Laney's work gives readers a deeper understanding of how and why information should be monetized and managed as an enterprise asset. Laney's assertion that accounting should recognize information as a capital asset is quite convincing and one I agree with. Infonomics enjoyably echoes that sentiment!" Matt Green, independent business analytics consultant, Atlanta area "If you care about the digital economy, and you should, read this book." Tanya Shuckhart, Analyst Relations Lead, IRI Worldwide

Creating a Data-Driven Organization - Carl Anderson 2015-07-23

"What do you need to become a data-driven organization? Far more than having big data or a crack team of unicorn data scientists, it requires establishing an effective, deeply-ingrained data culture. This practical book shows you how true data-drivenness involves processes that require genuine buy-in across your company ... Through

interviews and examples from data scientists and analytics leaders in a variety of industries ... Anderson explains the analytics value chain you need to adopt when building predictive business models"--Publisher's description.

Entity Resolution and Information Quality -

John R. Talburt 2011-01-14

Entity Resolution and Information Quality presents topics and definitions, and clarifies confusing terminologies regarding entity resolution and information quality. It takes a very wide view of IQ, including its six-domain framework and the skills formed by the International Association for Information and Data Quality (IAIDQ). The book includes chapters that cover the principles of entity resolution and the principles of Information Quality, in addition to their concepts and terminology. It also discusses the Fellegi-Sunter theory of record linkage, the Stanford Entity Resolution Framework, and the Algebraic Model for Entity Resolution, which are the major

theoretical models that support Entity Resolution. In relation to this, the book briefly discusses entity-based data integration (EBDI) and its model, which serve as an extension of the Algebraic Model for Entity Resolution. There is also an explanation of how the three commercial ER systems operate and a description of the non-commercial open-source system known as OYSTER. The book concludes by discussing trends in entity resolution research and practice. Students taking IT courses and IT professionals will find this book invaluable. First authoritative reference explaining entity resolution and how to use it effectively Provides practical system design advice to help you get a competitive advantage Includes a companion site with synthetic customer data for applicatory exercises, and access to a Java-based Entity Resolution program.

HBR Guide to Project Management (HBR Guide Series) - Harvard Business Review
2013-01-08

MEET YOUR GOALS—ON TIME AND ON BUDGET. How do you rein in the scope of your project when you've got a group of demanding stakeholders breathing down your neck? And map out a schedule everyone can stick to? And motivate team members who have competing demands on their time and attention? Whether you're managing your first project or just tired of improvising, this guide will give you the tools and confidence you need to define smart goals, meet them, and capture lessons learned so future projects go even more smoothly. The HBR Guide to Project Management will help you:

- Build a strong, focused team
- Break major objectives into manageable tasks
- Create a schedule that keeps all the moving parts under control
- Monitor progress toward your goals
- Manage stakeholders' expectations
- Wrap up your project and gauge its success

Execution - Larry Bossidy 2009-11-10
#1 NEW YORK TIMES BESTSELLER • More than two million copies in print! The premier

resource for how to deliver results in an uncertain world, whether you're running an entire company or in your first management job. "A must-read for anyone who cares about business."—The New York Times

When Execution was first published, it changed the way we did our jobs by focusing on the critical importance of "the discipline of execution": the ability to make the final leap to success by actually getting things done. Larry Bossidy and Ram Charan now reframe their empowering message for a world in which the old rules have been shattered, radical change is becoming routine, and the ability to execute is more important than ever. Now and for the foreseeable future:

- Growth will be slower. But the company that executes well will have the confidence, speed, and resources to move fast as new opportunities emerge.
- Competition will be fiercer, with companies searching for any possible advantage in every area from products and technologies to location and management.
-

Governments will take on new roles in their national economies, some as partners to business, others imposing constraints. Companies that execute well will be more attractive to government entities as partners and suppliers and better prepared to adapt to a new wave of regulation. • Risk management will become a top priority for every leader. Execution gives you an edge in detecting new internal and external threats and in weathering crises that can never be fully predicted. Execution shows how to link together people, strategy, and operations, the three core processes of every business. Leading these processes is the real job of running a business, not formulating a “vision” and leaving the work of carrying it out to others. Bossidy and Charan show the importance of being deeply and passionately engaged in an organization and why robust dialogues about people, strategy, and operations result in a business based on intellectual honesty and realism. With paradigmatic case histories from

the real world—including examples like the diverging paths taken by Jamie Dimon at JPMorgan Chase and Charles Prince at Citigroup—Execution provides the realistic and hard-nosed approach to business success that could come only from authors as accomplished and insightful as Bossidy and Charan.

Data Quality - Rupa Mahanti 2019-03-18

□This is not the kind of book that you□ll read one time and be done with. So scan it quickly the first time through to get an idea of its breadth. Then dig in on one topic of special importance to your work. Finally, use it as a reference to guide your next steps, learn details, and broaden your perspective.□ from the foreword by Thomas C. Redman, Ph.D., □the Data Doc□ Good data is a source of myriad opportunities, while bad data is a tremendous burden. Companies that manage their data effectively are able to achieve a competitive advantage in the marketplace, while bad data, like cancer, can weaken and kill an organization. In this comprehensive book, Rupa

Mahanti provides guidance on the different aspects of data quality with the aim to be able to improve data quality. Specifically, the book addresses: -Causes of bad data quality, bad data quality impacts, and importance of data quality to justify the case for data quality-Butterfly effect of data quality-A detailed description of data quality dimensions and their measurement-Data quality strategy approach-Six Sigma - DMAIC approach to data quality-Data quality management techniques-Data quality in relation to data initiatives like data migration, MDM, data governance, etc.-Data quality myths, challenges, and critical success factors Students, academicians, professionals, and researchers can all use the content in this book to further their knowledge and get guidance on their own specific projects. It balances technical details (for example, SQL statements, relational database components, data quality dimensions measurements) and higher-level qualitative discussions (cost of data quality, data quality

strategy, data quality maturity, the case made for data quality, and so on) with case studies, illustrations, and real-world examples throughout.

MASTER DATA MANAGEMENT AND DATA GOVERNANCE, 2/E - Alex Berson 2010-12-06

The latest techniques for building a customer-focused enterprise environment "The authors have appreciated that MDM is a complex multidimensional area, and have set out to cover each of these dimensions in sufficient detail to provide adequate practical guidance to anyone implementing MDM. While this necessarily makes the book rather long, it means that the authors achieve a comprehensive treatment of MDM that is lacking in previous works." -- Malcolm Chisholm, Ph.D., President, AskGet.com Consulting, Inc. Regain control of your master data and maintain a master-entity-centric enterprise data framework using the detailed information in this authoritative guide. Master Data Management and Data Governance, Second

Edition provides up-to-date coverage of the most current architecture and technology views and system development and management methods. Discover how to construct an MDM business case and roadmap, build accurate models, deploy data hubs, and implement layered security policies. Legacy system integration, cross-industry challenges, and regulatory compliance are also covered in this comprehensive volume. Plan and implement enterprise-scale MDM and Data Governance solutions Develop master data model Identify, match, and link master records for various domains through entity resolution Improve efficiency and maximize integration using SOA and Web services Ensure compliance with local, state, federal, and international regulations Handle security using authentication, authorization, roles, entitlements, and encryption Defend against identity theft, data compromise, spyware attack, and worm infection Synchronize components and test data quality

and system performance

The Saffron Kitchen - Yasmin Crowther

2007-08-28

In a powerful debut novel that moves between the crowded streets of London and the desolate mountains of Iran, Yasmin Crowther paints a stirring portrait of a family shaken by events from decades ago and worlds away. On a rainy day in London the dark secrets and troubled past of Maryam Mazar surface violently, with tragic consequences for her daughter, Sara, and her newly orphaned nephew. Maryam leaves her English husband and family and returns to the remote Iranian village where her story began. In a quest to piece their life back together, Sara follows her mother and finally learns the terrible price Maryam once had to pay for her freedom, and of the love she left behind. Set against the breathtaking beauty of two very different places, this stunning family drama transcends culture and is, at its core, a rich and haunting narrative about mothers and daughters.

Managing Data in Motion - April Reeve

2013-02-26

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"

Executing Data Quality Projects - Danette McGilvray 2021-05-27

Executing Data Quality Projects, Second Edition presents a structured yet flexible approach for creating, improving, sustaining and managing the quality of data and information within any organization. Studies show that data quality problems are costing businesses billions of dollars each year, with poor data linked to waste and inefficiency, damaged credibility among customers and suppliers, and an organizational inability to make sound decisions. Help is here! This book describes a proven Ten Step approach that combines a conceptual framework for understanding information quality with techniques, tools, and instructions for practically putting the approach to work - with the end result of high-quality trusted data and information, so critical to today's data-dependent organizations. The Ten Steps approach applies to all types of data and all types of organizations - for-profit in any

industry, non-profit, government, education, healthcare, science, research, and medicine. This book includes numerous templates, detailed examples, and practical advice for executing every step. At the same time, readers are advised on how to select relevant steps and apply them in different ways to best address the many situations they will face. The layout allows for quick reference with an easy-to-use format highlighting key concepts and definitions, important checkpoints, communication activities, best practices, and warnings. The experience of actual clients and users of the Ten Steps provide real examples of outputs for the steps plus highlighted, sidebar case studies called Ten Steps in Action. This book uses projects as the vehicle for data quality work and the word broadly to include: 1) focused data quality improvement projects, such as improving data used in supply chain management, 2) data quality activities in other projects such as building new applications and migrating data

from legacy systems, integrating data because of mergers and acquisitions, or untangling data due to organizational breakups, and 3) ad hoc use of data quality steps, techniques, or activities in the course of daily work. The Ten Steps approach can also be used to enrich an organization's standard SDLC (whether sequential or Agile) and it complements general improvement methodologies such as six sigma or lean. No two data quality projects are the same but the flexible nature of the Ten Steps means the methodology can be applied to all. The new Second Edition highlights topics such as artificial intelligence and machine learning, Internet of Things, security and privacy, analytics, legal and regulatory requirements, data science, big data, data lakes, and cloud computing, among others, to show their dependence on data and information and why data quality is more relevant and critical now than ever before. Includes concrete instructions, numerous templates, and practical advice for

executing every step of The Ten Steps approach. Contains real examples from around the world, gleaned from the author's consulting practice and from those who implemented based on her training courses and the earlier edition of the book. Allows for quick reference with an easy-to-use format 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 templates featured in the text, quick summaries of key ideas from the Ten Steps methodology, and other tools and information that are available online.

The Greenhouse Gas Protocol - World Resources Institute 2004-01-01

The GHG Protocol Corporate Accounting and Reporting Standard helps companies and other organizations to identify, calculate, and report GHG emissions. It is designed to set the standard for accurate, complete, consistent,

relevant and transparent accounting and reporting of GHG emissions.

Data Quality - Jack E. Olson 2003-01-09

Data Quality: The Accuracy Dimension is about assessing the quality of corporate data and improving its accuracy using the data profiling method. Corporate data is increasingly important as companies continue to find new ways to use it. Likewise, improving the accuracy of data in information systems is fast becoming a major goal as companies realize how much it affects their bottom line. Data profiling is a new technology that supports and enhances the accuracy of databases throughout major IT shops. Jack Olson explains data profiling and shows how it fits into the larger picture of data quality. * Provides an accessible, enjoyable introduction to the subject of data accuracy, peppered with real-world anecdotes. * Provides a framework for data profiling with a discussion of analytical tools appropriate for assessing data accuracy. * Is written by one of the original

developers of data profiling technology. * Is a must-read for any data management staff, IT management staff, and CIOs of companies with data assets.

Managing Data Quality - Tim King 2020-04-27

This book explains data quality management in practical terms, focusing on three key areas - the nature of data in enterprises, the purpose and scope of data quality management, and implementing a data quality management system, in line with ISO 8000-61. Examples of good practice in data quality management are also included.

Data Strategy - Sid Adelman 2005

Without a data strategy, the people within an organization have no guidelines for making decisions that are absolutely crucial to the success of the IT organization and to the entire organization. The absence of a strategy gives a blank check to those who want to pursue their own agendas, including those who want to try new database management systems, new

technologies (often unproven), and new tools. This type of environment provides no hope for success. Data Strategy should result in the development of systems with less risk, higher quality systems, and reusability of assets. This is key to keeping cost and maintenance down, thus running lean and mean. Data Strategy provides a CIO with a rationale to counter arguments for immature technology and data strategies that are inconsistent with existing strategies. This book uses case studies and best practices to give the reader the tools they need to create the best strategy for the organization.

Non-Invasive Data Governance - Robert S. Seiner 2014-09-01

Data-governance programs focus on authority and accountability for the management of data as a valued organizational asset. Data Governance should not be about command-and-control, yet at times could become invasive or threatening to the work, people and culture of an organization. Non-Invasive Data

Governance™ focuses on formalizing existing accountability for the management of data and improving formal communications, protection, and quality efforts through effective stewarding of data resources. Non-Invasive Data Governance will provide you with a complete set of tools to help you deliver a successful data governance program. Learn how:

- Steward responsibilities can be identified and recognized, formalized, and engaged according to their existing responsibility rather than being assigned or handed to people as more work.
- Governance of information can be applied to existing policies, standard operating procedures, practices, and methodologies, rather than being introduced or emphasized as new processes or methods.
- Governance of information can support all data integration, risk management, business intelligence and master data management activities rather than imposing inconsistent rigor to these initiatives.
- A practical and non-threatening approach can be

applied to governing information and promoting stewardship of data as a cross-organization asset. • Best practices and key concepts of this non-threatening approach can be communicated effectively to leverage strengths and address opportunities to improve.

Multi-Domain Master Data Management - Mark Allen 2015-03-21

Multi-Domain Master Data Management delivers practical guidance and specific instruction to help guide planners and practitioners through the challenges of a multi-domain master data management (MDM) implementation. Authors Mark Allen and Dalton Cervo bring their expertise to you in the only reference you need to help your organization take master data management to the next level by incorporating it across multiple domains. Written in a business friendly style with sufficient program planning guidance, this book covers a comprehensive set of topics and advanced strategies centered on the key MDM disciplines of Data Governance,

Data Stewardship, Data Quality Management, Metadata Management, and Data Integration. Provides a logical order toward planning, implementation, and ongoing management of multi-domain MDM from a program manager and data steward perspective. Provides detailed guidance, examples and illustrations for MDM practitioners to apply these insights to their strategies, plans, and processes. Covers advanced MDM strategy and instruction aimed at improving data quality management, lowering data maintenance costs, and reducing corporate risks by applying consistent enterprise-wide practices for the management and control of master data.

A Guide to the Project Management Body of Knowledge (PMBOK® Guide) - Seventh Edition and The Standard for Project Management (BRAZILIAN PORTUGUESE) - Project Management Institute Project Management Institute 2021-08-01
PMBOK® Guide is the go-to resource for

project management practitioners. The project management profession has significantly evolved due to emerging technology, new approaches and rapid market changes. Reflecting this evolution, The Standard for Project Management enumerates 12 principles of project management and the PMBOK® Guide – Seventh Edition is structured around eight project performance domains. This edition is designed to address practitioners' current and future needs and to help them be more proactive, innovative and nimble in enabling desired project outcomes. This edition of the PMBOK® Guide:

- Reflects the full range of development approaches (predictive, adaptive, hybrid, etc.);
- Provides an entire section devoted to tailoring the development approach and processes;
- Includes an expanded list of models, methods, and artifacts;
- Focuses on not just delivering project outputs but also enabling outcomes; and
- Integrates with PMI standards+™ for information and standards

application content based on project type, development approach, and industry sector.

Handbook of Data Quality - Shazia Sadiq
2013-08-13

The issue of data quality is as old as data itself. However, the proliferation of diverse, large-scale and often publically available data on the Web has increased the risk of poor data quality and misleading data interpretations. On the other hand, data is now exposed at a much more strategic level e.g. through business intelligence systems, increasing manifold the stakes involved for individuals, corporations as well as government agencies. There, the lack of knowledge about data accuracy, currency or completeness can have erroneous and even catastrophic results. With these changes, traditional approaches to data management in general, and data quality control specifically, are challenged. There is an evident need to incorporate data quality considerations into the whole data cycle, encompassing

managerial/governance as well as technical aspects. Data quality experts from research and industry agree that a unified framework for data quality management should bring together organizational, architectural and computational approaches. Accordingly, Sadiq structured this handbook in four parts: Part I is on organizational solutions, i.e. the development of data quality objectives for the organization, and the development of strategies to establish roles, processes, policies, and standards required to manage and ensure data quality. Part II, on architectural solutions, covers the technology landscape required to deploy developed data quality management processes, standards and policies. Part III, on computational solutions, presents effective and efficient tools and techniques related to record linkage, lineage and provenance, data uncertainty, and advanced integrity constraints. Finally, Part IV is devoted to case studies of successful data quality initiatives that highlight the various aspects of

data quality in action. The individual chapters present both an overview of the respective topic in terms of historical research and/or practice and state of the art, as well as specific techniques, methodologies and frameworks developed by the individual contributors. Researchers and students of computer science, information systems, or business management as well as data professionals and practitioners will benefit most from this handbook by not only focusing on the various sections relevant to their research area or particular practical work, but by also studying chapters that they may initially consider not to be directly relevant to them, as there they will learn about new perspectives and approaches.

Data Quality - Carlo Batini 2006-09-27

Poor data quality can seriously hinder or damage the efficiency and effectiveness of organizations and businesses. The growing awareness of such repercussions has led to major public initiatives like the "Data Quality Act" in the USA and the

"European 2003/98" directive of the European Parliament. Batini and Scannapieco present a comprehensive and systematic introduction to the wide set of issues related to data quality. They start with a detailed description of different data quality dimensions, like accuracy, completeness, and consistency, and their importance in different types of data, like federated data, web data, or time-dependent data, and in different data categories classified according to frequency of change, like stable, long-term, and frequently changing data. The book's extensive description of techniques and methodologies from core data quality research as well as from related fields like data mining, probability theory, statistical data analysis, and machine learning gives an excellent overview of the current state of the art. The presentation is completed by a short description and critical comparison of tools and practical methodologies, which will help readers to resolve their own quality problems. This book is an ideal

combination of the soundness of theoretical foundations and the applicability of practical approaches. It is ideally suited for everyone – researchers, students, or professionals – interested in a comprehensive overview of data quality issues. In addition, it will serve as the basis for an introductory course or for self-study on this topic.

Subsystem and Transaction Monitoring and Tuning with DB2 11 for z/OS - Paolo Bruni
2015-07-29

This IBM® Redbooks® publication discusses in detail the facilities of DB2® for z/OS®, which allow complete monitoring of a DB2 environment. It focuses on the use of the DB2 instrumentation facility component (IFC) to provide monitoring of DB2 data and events and includes suggestions for related tuning. We discuss the collection of statistics for the verification of performance of the various components of the DB2 system and accounting for tracking the behavior of the applications. We

have intentionally omitted considerations for query optimization; they are worth a separate document. Use this book to activate the right traces to help you monitor the performance of your DB2 system and to tune the various aspects of subsystem and application performance.

Executing Data Quality Projects - Danette McGilvray 2008-09-01

Information is currency. Recent studies show that data quality problems are costing businesses billions of dollars each year, with poor data linked to waste and inefficiency, damaged credibility among customers and suppliers, and an organizational inability to make sound decisions. In this important and timely new book, Danette McGilvray presents her "Ten Steps approach to information quality, a proven method for both understanding and creating information quality in the enterprise. Her trademarked approach—in which she has trained Fortune 500 clients and hundreds of workshop attendees—applies to all types of data

and to all types of organizations. * Includes numerous templates, detailed examples, and practical advice for executing every step of the "Ten Steps approach. * Allows for quick reference with an easy-to-use format 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.

Learning How to Learn - Barbara Oakley, PhD 2018-08-07

A surprisingly simple way for students to master any subject--based on one of the world's most popular online courses and the bestselling book A Mind for Numbers A Mind for Numbers and its wildly popular online companion course "Learning How to Learn" have empowered more

than two million learners of all ages from around the world to master subjects that they once struggled with. Fans often wish they'd discovered these learning strategies earlier and ask how they can help their kids master these skills as well. Now in this new book for kids and teens, the authors reveal how to make the most of time spent studying. We all have the tools to learn what might not seem to come naturally to us at first--the secret is to understand how the brain works so we can unlock its power. This book explains:

- Why sometimes letting your mind wander is an important part of the learning process
- How to avoid "rut think" in order to think outside the box
- Why having a poor memory can be a good thing
- The value of metaphors in developing understanding
- A simple, yet powerful, way to stop procrastinating

Filled with illustrations, application questions, and exercises, this book makes learning easy and fun.

Registries for Evaluating Patient Outcomes -

Agency for Healthcare Research and Quality/AHRQ 2014-04-01

This User's Guide is intended to support the design, implementation, analysis, interpretation, and quality evaluation of registries created to increase understanding of patient outcomes. For the purposes of this guide, a patient registry is an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical, or policy purposes. A registry database is a file (or files) derived from the registry. Although registries can serve many purposes, this guide focuses on registries created for one or more of the following purposes: to describe the natural history of disease, to determine clinical effectiveness or cost-effectiveness of health care products and services, to measure or monitor safety and harm, and/or to measure quality of

care. Registries are classified according to how their populations are defined. For example, product registries include patients who have been exposed to biopharmaceutical products or medical devices. Health services registries consist of patients who have had a common procedure, clinical encounter, or hospitalization. Disease or condition registries are defined by patients having the same diagnosis, such as cystic fibrosis or heart failure. The User's Guide was created by researchers affiliated with AHRQ's Effective Health Care Program, particularly those who participated in AHRQ's DEcIDE (Developing Evidence to Inform Decisions About Effectiveness) program. Chapters were subject to multiple internal and external independent reviews.

Graph Algorithms - Mark Needham 2019-05-16

Discover how graph algorithms can help you leverage the relationships within your data to develop more intelligent solutions and enhance your machine learning models. You'll learn how

graph analytics are uniquely suited to unfold complex structures and reveal difficult-to-find patterns lurking in your data. Whether you are trying to build dynamic network models or forecast real-world behavior, this book illustrates how graph algorithms deliver value—from finding vulnerabilities and bottlenecks to detecting communities and improving machine learning predictions. This practical book walks you through hands-on examples of how to use graph algorithms in Apache Spark and Neo4j—two of the most common choices for graph analytics. Also included: sample code and tips for over 20 practical graph algorithms that cover optimal pathfinding, importance through centrality, and community detection. Learn how graph analytics vary from conventional statistical analysis. Understand how classic graph algorithms work, and how they are applied. Get guidance on which algorithms to use for different types of questions. Explore algorithm examples with working code.

and sample datasets from Spark and Neo4j See how connected feature extraction can increase

machine learning accuracy and precision Walk through creating an ML workflow for link prediction combining Neo4j and Spark