

Data Warehouse From Architecture To Implementation

In this book, I introduce a new architectural layer that integrates SOA applications with a near-real-time data warehouse. Through this new "SOA Integration Layer" a new Information Technology (IT) solution is created that realizes even greater benefits to the business as a result of combining these individual information technologies. This new IT solution will connect the SOA integration layer to near-real-time data warehouse and consider the near-real-time data warehouse as a data stage source for other business and system environments. This new IT solution is unique because it focuses on relating SOA application environments to near-real-time data warehouse to support both business operations and business strategic decision making. This study will propose a new prototype model in order to build a SOA integration layer and present the SOA integration layer and to connect this SOA integration layer to the data management layer which includes a near-real-time data warehouse. The expected outcomes from this book will be combining near-real-time data warehousing with SOA architecture to deliver integrated tactical and strategic data. Organization of data warehouses is a vital, but often neglected, aspect of growing an enterprise. Unlike most books on the subject that focus on either the technical aspects of building data warehouses or on business strategies, this valuable reference synthesizes technological know-how with managerial best practices to show how improved alignment between data warehouse plans and business strategies can lead to successful data warehouse adoption capable of supporting an enterprise's entire infrastructure. Strategic Data Warehousing: Achieving Alignment with Business provides data warehouse developers, business managers, and IT professionals and administrators with an integrated approach to achieving successful and sustainable alignment of data warehouses and business goals. More complete than any other text in the field, this comprehensive reference details the joint roles and responsibilities of the data warehouse and business managers in achieving strategic alignment, business user satisfaction, technical integration, and improved flexibility. Complete with case studies that depict real-world scenarios, the text: Examines the organizational, user, data, and technological factors proven to promote successful data warehousing Includes actionable solutions for achieving strategic alignment Provides a model that readers can apply in aligning their own data warehouse needs and business goals Achieving sustainable alignment between the data warehouse and business strategies is a continuous process. Armed with this valuable reference, readers will be able to gain the solid understanding of the organizational, technical, data, and user factors needed to promote a successful data warehouse adoption and become active partners in leveraging this powerful, but often overlooked, information resource.

In this all-in-one, Tourism-oriented Data Warehousing architecture is proposed which makes use of the most recent free and open source technologies like Java, PostgreSQL and XML. Such architecture's aim will be to support the decision making process and giving an integrated view of the whole Tourism realm in an established context (local, regional, national, etc.) without requesting big investments for getting the necessary software.

The first, step-by-step guide to building Web-enabled data warehouses The Web can be an incredibly rich source of customer data, and right now companies across industry sectors are hustling to get up and running with data warehouses capable of capturing the clickstream data from their Web sites. This allows companies to track exactly where a customer is going, or "clicking to," on their site in order to gain meaningful information about that customer's preferences. Following Ralph Kimball's The Data Warehouse Toolkit (0-717-37680-9) where he provides the blueprint, Clickstream Data Warehousing fills developers in on all the technical details that go into building a Web-enabled data warehouse. The authors review all key architectural and design issues that developers need to masterfully build a Webhouse using examples to illustrate key points. Companion Web site features code examples from the book and links to related Web sites.

A Guide for Solution Architects and Project Leaders

Snowflake Cookbook

Learn Data Warehousing in 24 Hours

Building a Data Warehouse for Decision Support

Architecture and Implementation

Data Warehousing Fundamentals

The new edition of the classic bestseller that launched the data warehousing industry covers new approaches and technologies, many of which have been pioneered by Immon himself. In addition to explaining the fundamentals of data warehouse systems, the book covers new topics such as methods for handling unstructured data in a data warehouse, accessing data across multiple stage media, accessing the pros and cons of relational versus multidimensional design and how to measure return on investment in planning data warehouse projects. Covers advanced topics, including data monitoring and testing. Although the book includes an extra 100 pages worth of valuable content, the price has actually been reduced from \$65 to \$55.

A thorough update to the industry standard for designing, developing, and deploying data warehouse and business intelligence systems. The world of data warehousing has changed remarkably since the first edition of The Data Warehouse Lifecycle Toolkit was published in 1998. In that time, the data warehouse industry has reached full maturity and acceptance, hardware and software have made staggering advances, and the techniques promoted in the premiere edition of this book have been adopted by nearly all data warehouse vendors and practitioners. In addition, the term "business intelligence" emerged to reflect the mission of the data warehouse: wrangling the data out of source systems, cleaning it, and delivering it to add value to the business. Ralph Kimball and his colleagues have refined the original set of Lifecycle methods and techniques based on their consulting and training experience. The authors understand first-hand that a data warehousing/business intelligence (DW/BI) system needs to change as fast as its surrounding organization evolves. To that end, they walk you through the detailed steps of designing, developing, and deploying a DW/BI system. You'll learn to create adaptable systems that deliver data and analyses to business users so they can make better business decisions.

The "father of data warehousing" incorporates the latest technologies into his blueprint for integrated decision support systems. Today's corporate IT and data warehouse managers are required to make a small army of technologies work together to ensure fast and accurate information for business managers. Bill Immon created the Corporate Information Factory to solve the needs of these managers. Since the First Edition, the design of the factory has grown and changed dramatically. This Second Edition, revised and expanded by 40% with five new chapters, incorporates these changes. This step-by-step guide will enable readers to connect their legacy systems with the data warehouse and deal with a host of new and changing technologies, including Web access mechanisms, e-commerce systems, ERP (Enterprise Resource Planning) systems. The book also looks closely at exploration and data mining servers for analyzing customer behavior and departmental data marts for finance, sales, and marketing.

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 of 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."

A Primer for the Data Scientist

Data Warehousing in the Age of Big Data

Clickstream Data Warehousing

Building the Unstructured Data Warehouse

Building a Scalable Data Warehouse with Data Vault 2.0

Building a Data Warehouse: With Examples in SQL Server describes how to build a data warehouse completely from scratch and shows practical examples on how to do it. Author Vincent Rainardi also describes some practical issues he has experienced that developers are likely to encounter in their first data warehousing project, along with database management system (RDBMS) used in the examples is SQL Server. The version will not be an issue as long as the user has SQL Server 2005 or later. The book is organized as follows. In the beginning of this book (chapters 1 through 6), you learn how to build a data warehouse, for example, defining the architecture, understanding requirements, designing the data models, and creating the databases. Then in chapters 7 through 10, you learn how to populate the data warehouse, for example, extracting from source systems, loading the data stores, maintaining data quality, and utilizing the metadata. After you populate the data warehouse, in chapters 11 through 15, reports and multidimensional databases and how to use the data in the data warehouse for business intelligence, customer relationship management, and other purposes. Chapters 16 and 17 wrap up the book. After you have built your data warehouse, before it can be released to production, you need to test it thoroughly. After your work, you understand how to administer data warehouse operation. What you'll learn A detailed understanding of what it takes to build a data warehouse. The implementation code in SQL Server to build the data warehouse. Dimensional modeling, data extraction methods, data warehouse loading, populating dimension and fact tables, data quality, data design. Practical data warehousing applications such as business intelligence reports, analytics applications, and customer relationship management. Who this book is for There are three audiences for the book. The first are the people who implement the data warehouse. This could be considered a field guide for them. The second is database administrators who want to build a data warehouse. Finally, the third audience is managers who must make decisions about aspects of the data warehousing task before them and use the book to learn about these issues.

Organizations invest incredible amounts of time and money obtaining and then storing big data in data stores called data lakes. But how many of these organizations can actually get the data back out in a useable form? Very few can turn the data lake into an information gold mine. Most wind up with garbage dumps. Data Lake Architecture

lake, where data scientists and data analysts can solve business challenges and identify new business opportunities. Learn how to structure data lakes as well as analog application, and text-based data ponds to provide maximum business value. Understand the role of the raw data pond and when to use an archival data pond. Leverage the

metadata, integration mapping, context, and metaprocess. Bill Immon opened our eyes to the architecture and benefits of a data warehouse, and now he takes us to the next level of data lake architecture.

The Data Vault was invented by Dan Linstedt at the U.S. Department of Defense, and the standard has been successfully applied to data warehousing projects at organizations of different sizes, from small to large-size corporations. Due to its simplified design, which is adapted from nature, the Data Vault 2.0 standard helps prevent typical Scatible Data Warehouse" covers everything one needs to know to create a scalable data warehouse end to end, including a presentation of the Data Vault modeling technique, which provides the foundations to create a technical data warehouse layer. The book discusses how to build the data warehouse incrementally using the agile Data Vault will learn how to create the input layer (the stage layer) and the presentation layer (data mart) of the Data Vault 2.0 architecture including implementation best practices. Drawing upon years of practical experience and using numerous examples and an easy to understand framework, Dan Linstedt and Michael Olschmke discuss: How to leverage Services (SSIS), including automation of the Data Vault loading processes. Important data warehouse technologies and practices. Data Quality Services (DQS) and Master Data Services (MDS) in the context of the Data Vault architecture. Provides a complete introduction to data warehousing, applications, and the business context so readers can understand theoretical concepts and provides hands-on instruction on how to build and implement a data warehouse. Demystifies data vault modeling with beginning, intermediate, and advanced techniques. Discusses the advantages of the data vault approach over other techniques, also including the latest updates to Data Vault 2.0 and multiple improve

The data lakehouse is the next generation of the data warehouse and data lake, designed to meet today's complex and ever-changing analytics, machine learning, and data science requirements. Learn about the features and architecture of the data lakehouse, along with its powerful analytical infrastructure. Appreciate how the universal cloud analog, and IoT data. Maintain the lakehouse for future generations through Data Lakehouse Housekeeping and Data Future-proofing. Know how to incorporate the lakehouse into an existing data governance strategy. Incorporate data catalogs, data lineage tools, and open source software into your architecture to ensure your data scientists

after.

Managing Data in Motion

Data Warehousing with Service-Oriented Architecture

Agile Data Warehousing for the Enterprise

Data Integration Best Practice Techniques and Technologies

From Architecture to Implementation

Corporate Information Factory

Data warehousing is one of the hottest business topics, and there's more to understanding data warehousing technologies than you might think. Find out the basics of data warehousing and how it facilitates data mining and business intelligence with Data Warehousing For Dummies, 2nd Edition. Data is probably your company's most important asset, so your data warehouse should serve your needs. The fully updated Second Edition of Data Warehousing For Dummies helps you understand, develop, implement, and use data warehouses, and offers a sneak peek into their future. You'll learn to: Analyze top-down and bottom-up data warehouse designs. Understand the structure and technologies of data warehouses, operational data stores, and data marts. Choose your project team and apply best development practices to your data warehousing projects. Implement a data warehouse, step by step, and involve end-users in the process. Review and upgrade existing data storage to make it serve your needs. Comprehend OLAP, column-wise databases, hardware-assisted databases, and middleware. Use data mining intelligently and find what you need. Make informed choices between consultants and data warehousing products. Data Warehousing For Dummies, 2nd Edition also shows you how to involve users in the testing process and gain valuable feedback, what it takes to successfully manage a data warehouse project, and how to tell if your project is on track. You'll find it's the most useful source of data on the topic!

DW 2.0: The Next Generation of Data Warehousing is the first book on the new generation of data warehouse architecture, DW 2.0, by the father of the data warehouse. The book describes the future of data warehousing that is technologically possible today, at both an architectural level and technology level. The perspective of the book is from the top down: looking at the overall architecture and then delving into the issues underlying the components. This allows people who are building or using a data warehouse to see what lies ahead and determine what new technology to buy, how to plan extensions to the data warehouse, what can be salvaged from the current system, and how to justify the expense at the most practical level. This book gives experienced data warehouse professionals everything they need in order to implement the new generation DW 2.0. It is designed for professionals in the IT organization, including data architects, DBAs, systems design and development professionals, as well as data warehouse and knowledge management professionals. * First book on the new generation of data warehouse architecture, DW 2.0. * Written by the "father of the data warehouse," Bill Immon, a columnist and newsletter editor of The Bill Immon Channel on the Business Intelligence Network. * Long overdue comprehensive coverage of the implementation of technology and tools that enable the new generation of the DW: metadata, temporal data, ETL, unstructured data, and data quality control.

Here is the ideal field guide for data warehousing implementation. This book first teaches you how to build a data warehouse, including defining the architecture, understanding the methodology, gathering the requirements, designing the data models, and creating the databases. Coverage then explains how to populate the data warehouse and explores how to present data to users using reports and multidimensional databases and how to use the data in the data warehouse for business intelligence, customer relationship management, and other purposes. It also details testing and how to administer data warehouse operation.

Building upon his earlier book that detailed agile data warehousing programming techniques for the Scrum master, Ralph's latest work illustrates the agile interpretations of the remaining software engineering disciplines: Requirements management benefits from streamlining templates that not only define projects quickly, but ensure nothing essential is overlooked. Data engineering receives two new "hybrid modeling" techniques, yielding data warehouses that can be easily adapted when requirements change without having to invest in ruinously expensive data-conversion programs. Quality assurance advances with not only a stereoscopic top-down and bottom-up planning method, but also the incorporation of the latest in automated test engines. Use this step-by-step guide to deepen your own application development skills through self-study, show your teammates the world's fastest and most reliable techniques for creating business intelligence systems, or ensure that the IT department working for you is building your next decision support system the right way. Learn how to quickly define scope and architecture before programming starts. Includes techniques of process and data engineering that enable iterative and incremental delivery. Demonstrates how to plan and execute quality assurance plans and includes a guide to continuous integration and automated regression testing. Presents program management strategies for coordinating multiple agile data mart projects so that over time an enterprise data warehouse emerges. Use the provided 120-day road map to establish a robust, agile data warehousing program.

Achieving Alignment with Business

Data Lake Architecture

With Examples in SQL Server

Super Charge Your Data Warehouse

Data Warehousing for E-Business

Big Data, Data Warehouse and Data Vault

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 appendices 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.

Do You Know How Your Data Warehouse Flexible, Scalable, Secure and Will It Stand The Test Of Time And Avoid Being Part Of The Dreaded "Life Cycle"? The Data Vault took the Data Warehouse world by storm when it was released in 2001. Some of the world's largest and most complex data warehouse situations understood the value it gave especially with the capabilities of end scaling, flexibility and security. Here is what industry leaders say about the Data Vault: "The Data Vault is the optimal choice for modeling the EDW in the DW 2.0 framework" - Bill Immon, The Father of Data Warehousing "The Data Vault is foundationally strong and an exceptionally scalable architecture" - Stephen Brobst, CTO, Teradata "The Data Vault should be considered as a potential standard for RDBMS-based analytic data management by organizations looking to achieve a high degree of flexibility, performance and openness" - Doug Laney, Deloitte Analytics Institute "I applaud Dan's contribution to the body of Business Intelligence and Data Warehousing knowledge and recommend this book be read by both data professionals and end users." - Howard Dresner, From the Foreword - Speaker, Author, Leading Research Analyst and Advisor You have in your hands the work, experience and testing of 2 decades of building data warehouses. The Data Vault model and methodology has proven itself in hundreds (perhaps thousands) of solutions in Insurance, Crime-Fighting, Defense, Retail, Finance, Banking, Power, Energy, Education, High-Tech and many more. Learn the techniques and implement them and learn how to build your Data Warehouse faster than you have ever done before while designing it to grow and scale no matter what you throw at it. Ready to "Super Charge Your Data Warehouse"?

Over the past 5 years, the concept of big data has matured, data science has grown exponentially, and data architecture has become a standard part of organizational decision-making. Throughout all this change, the basic principles that shape the architecture of data have remained the same. There remains a need for people to take a look at the "bigger picture" and to understand where their data fit into the grand scheme of things. Data Architecture: A Primer for the Data Scientist, Second Edition addresses the larger architectural picture of how big data fits within the existing information infrastructure or data warehousing systems. This is an essential topic not only for data scientists, analysts, and managers but also for researchers and engineers who increasingly need to deal with large and complex sets of data. Until data are gathered and can be placed into an existing framework or architecture, they cannot be used to their full potential. Drawing upon years of practical experience and using numerous examples and case studies from across various industries, the authors seek to explain this larger picture into which big data fits, giving data scientists the necessary context for how pieces of the puzzle should fit together. New case studies include expanded coverage of textual management and analytics. New chapters on visualization and big data. Discussion of new visualizations of the end-state architecture.

Today, the world is trying to create and educate data scientists because of the phenomenon of Big Data. And everyone is looking deeply into this technology. But no one is looking at the larger architectural picture of how Big Data needs to fit within the existing systems (data warehousing systems). Taking a look at the larger picture into which Big Data fits gives the data scientist the necessary context for how pieces of the puzzle should fit together. Most references on Big Data look at only one tiny part of a much larger whole. Until data gathered can be put into an existing framework or architecture it can't be used to its full potential. Data Architecture: A Primer for the Data Scientist addresses the larger architectural picture of how Big Data fits with the existing information infrastructure, an essential topic for the data scientist. Drawing upon years of practical experience and using numerous examples and an easy to understand framework, W.H. Immon, and Daniel Linstedt define the importance of data architecture and how it can be used effectively to harness big data within existing systems. You'll be able to: Turn textual information into a form that can be analyzed by standard tools. Make the connection between Analytics and Big Data. Understand how Big Data fits within an existing systems environment. Conduct analytics on repetitive and non-repetitive data. Discusses the value in Big Data that is often overlooked, non-repetitive data, and why there is significant business value in using it. Shows how to turn textual information into a form that can be analyzed by standard tools. Explains how Big Data fits within an existing systems environment. Presents new opportunities that are afforded by the advent of Big Data. Demystifies the murky waters of repetitive and non-repetitive data in Big Data.

The Data Warehouse Lifecycle Toolkit

with IBM Business Intelligence Tools

Building the Data Warehouse

The Data Warehouse Toolkit

DW 2.0: The Architecture for the Next Generation of Data Warehousing

IBM Data Warehousing

Data warehousing is one of the hottest topics in the computing industry. Written by Barry Devlin, one of the world's leading experts on data warehousing, this book gives you the insights and experiences gained over 10 years and offers the most comprehensive, practical guide to designing, building, and implementing a successful data warehouse. Included in this vital information is an explanation of the optimal three-tiered architecture for the data warehouse, with a clear division between data and information. Information systems managers will appreciate the full description of the functions needed to implement such an architecture, including metadata and deriving consistent, valuable business information.

As huge volumes of data, 24/7 performance requirements, and the need to integrate customer information are outstripping the capacity of many existing data warehouses, the question of how to support this explosive growth of e-business is becoming the hottest topic in data warehousing. Building upon Bill Immon's proven corporate information factory (CIF) architecture, this team of experts show data warehouse managers and developers how to architect the data warehouse environment to meet the high-performance requirements of data-intensive Web sites as well as the needs of users across the enterprise. Readers will get detailed guidance on how to build the data warehouse or improve existing capabilities in order to: o Capture clickstream data that customers enter at the Webhouse o Interface with HTML, XML, and other Web applications o Integrate clickstream data with legacy database systems o Design the data warehouse for optimum performance and scalability o Store and archive high volumes of data at reasonable costs o Analyze e-business data to spot new marketing and sales opportunities

Building upon their data warehousing, data warehouse architecture and infrastructure, and the various methods for information delivery. This practical Second Edition highlights the areas of data warehousing and business intelligence where high-impact technological progress has been made. Discussions on developments include data marts, real-time information delivery, data visualization, and "integrated" and "near-real-time." However, the organization's ability to successfully exploit their data is determined by the degree of strategic alignment. As such, this study poses the question: how can a data warehouse be successfully and demonstrably aligned to an organization's strategic objectives? This thesis demonstrates that strategic alignment can be achieved by following a new "top down" data warehouse implementation framework, the Configuration Approach, which is based upon determining an organization's target configuration. This was achieved by employing Miles and Snow's Ideal Types to formulate a questionnaire that reveals an organization's target configuration in terms of its approach to the Entrepreneurial, Administration and Information Systems challenges. Crucially, this thesis also provides the means to choose a data warehouse architecture that is wholly based on the organization's target configuration. The Configuration Approach was evaluated using a single case study undergoing a period of strategic transformation where the implementation of a data warehouse was key to its strategic ambitions. The case study illustrated how it is possible to articulate an organization's strategic configuration, which becomes the key driver for building a warehouse that demonstrably supports the resolution of its Entrepreneurial and Administration challenges. Significantly, the case study also provides a unique opportunity to demonstrate how the target configuration helps organizations to make the right choice of data warehouse architecture to satisfy the Information Systems challenge. In this case, the Configuration Approach provides a basis for challenging the architectural choices made by a consultancy on behalf of the participating organization. Accordingly, it can be asserted that data warehouses are strategic investments, if implemented using the Configuration Approach.

This book presents the first comparative review of the state of the art and the best current practices of data warehouses. It covers source and data integration, multidimensional aggregation, query optimization, metadata management, quality assessment, and design optimization. A conceptual framework is presented by which the architecture and quality of a data warehouse can be assessed and improved using enriched metadata management combined with advanced techniques from databases, business modeling, and artificial intelligence.

Data Warehousing

An Open Source Based Data Warehouse Architecture to Support Decision Making in the Tourism Sector

Strategic Data Warehousing

Building a Star Schema

Architecture, Analysis, and Design

Data Warehousing in Action

CUTTING-EDGE CONTENT AND GUIDANCE FROM A DATA WAREHOUSING EXPERT—NOW EXPANDED TO REFLECT FIELD TRENDS Data warehousing has revolutionized the way businesses in a wide variety of industries perform analysis and make strategic decisions. Since the first edition of Data Warehousing Fundamentals, numerous enterprises have implemented data warehouse systems and reaped the benefits. Now, this new, revised edition covers the essential fundamentals of data warehousing and business intelligence as well as significant recent trends in the field. The author provides an enhanced, comprehensive overview of data warehousing together with in-depth explanations of critical issues in planning, design, deployment, and ongoing maintenance. IT professionals eager to get into the field of modern data warehousing, data warehouse architecture and infrastructure, and the various methods for information delivery. This practical Second Edition highlights the areas of data warehousing and business intelligence where high-impact technological progress has been made. Discussions on developments include data marts, real-time information delivery, data visualization, and "integrated" and "near-real-time." However, the organization's ability to successfully exploit their data is determined by the degree of strategic alignment. As such, this study poses the question: how can a data warehouse be successfully and demonstrably aligned to an organization's strategic objectives? This thesis demonstrates that strategic alignment can be achieved by following a new "top down" data warehouse implementation framework, the Configuration Approach, which is based upon determining an organization's target configuration. This was achieved by employing Miles and Snow's Ideal Types to formulate a questionnaire that reveals an organization's target configuration in terms of its approach to the Entrepreneurial, Administration and Information Systems challenges. Crucially, this thesis also provides the means to choose a data warehouse architecture that is wholly based on the organization's target configuration. The Configuration Approach was evaluated using a single case study undergoing a period of strategic transformation where the implementation of a data warehouse was key to its strategic ambitions. The case study illustrated how it is possible to articulate an organization's strategic configuration, which becomes the key driver for building a warehouse that demonstrably supports the resolution of its Entrepreneurial and Administration challenges. Significantly, the case study also provides a unique opportunity to demonstrate how the target configuration helps organizations to make the right choice of data warehouse architecture to satisfy the Information Systems challenge. In this case, the Configuration Approach provides a basis for challenging the architectural choices made by a consultancy on behalf of the participating organization. Accordingly, it can be asserted that data warehouses are strategic investments, if implemented using the Configuration Approach.

This book presents the first comparative review of the state of the art and the best current practices of data warehouses. It covers source and data integration, multidimensional aggregation, query optimization, metadata management, quality assessment, and design optimization. A conceptual framework is presented by which the architecture and quality of a data warehouse can be assessed and improved using enriched metadata management combined with advanced techniques from databases, business modeling, and artificial intelligence.

An Open Source Based Data Warehouse Architecture to Support Decision Making in the Tourism Sector

Strategic Data Warehousing

Building a Star Schema

Architecture, Analysis, and Design

Data Warehousing in Action

CUTTING-EDGE CONTENT AND GUIDANCE FROM A DATA WAREHOUSING EXPERT—NOW EXPANDED TO REFLECT FIELD TRENDS Data warehousing has revolutionized the way businesses in a wide variety of industries perform analysis and make strategic decisions. Since the first edition of Data Warehousing Fundamentals, numerous enterprises have implemented data warehouse systems and reaped the benefits. Now, this new, revised edition covers the essential fundamentals of data warehousing and business intelligence as well as significant recent trends in the field. The author provides an enhanced, comprehensive overview of data warehousing together with in-depth explanations of critical issues in planning, design, deployment, and ongoing maintenance. IT professionals eager to get into the field of modern data warehousing, data warehouse architecture and infrastructure, and the various methods for information delivery. This practical Second Edition highlights the areas of data warehousing and business intelligence where high-impact technological progress has been made. Discussions on developments include data marts, real-time information delivery, data visualization, and "integrated" and "near-real-time." However, the organization's ability to successfully exploit their data is determined by the degree of strategic alignment. As such, this study poses the question: how can a data warehouse be successfully and demonstrably aligned to an organization's strategic objectives? This thesis demonstrates that strategic alignment can be achieved by following a new "top down" data warehouse implementation framework, the Configuration Approach, which is based upon determining an organization's target configuration. This was achieved by employing Miles and Snow's Ideal Types to formulate a questionnaire that reveals an organization's target configuration in terms of its approach to the Entrepreneurial, Administration and Information Systems challenges. Crucially, this thesis also provides the means to choose a data warehouse architecture that is wholly based on the organization's target configuration. The Configuration Approach was evaluated using a single case study undergoing a period of strategic transformation where the implementation of a data warehouse was key to its strategic ambitions. The case study illustrated how it is possible to articulate an organization's strategic configuration, which becomes the key driver for building a warehouse that demonstrably supports the resolution of its Entrepreneurial and Administration challenges. Significantly, the case study also provides a unique opportunity to demonstrate how the target configuration helps organizations to make the right choice of data warehouse architecture to satisfy the Information Systems challenge. In this case, the Configuration Approach provides a basis for challenging the architectural choices made by a consultancy on behalf of the participating organization. Accordingly, it can be asserted that data warehouses are strategic investments, if implemented using the Configuration Approach.

This book presents the first comparative review of the state of the art and the best current practices of data warehouses. It covers source and data integration, multidimensional aggregation, query optimization, metadata management, quality assessment, and design optimization. A conceptual framework is presented by which the architecture and quality of a data warehouse can be assessed and improved using enriched metadata management combined with advanced techniques from databases, business modeling, and artificial intelligence.

An Open Source Based Data Warehouse Architecture to Support Decision Making in the Tourism Sector

Strategic Data Warehousing

Building a Star Schema

Architecture, Analysis, and Design

Data Warehousing in Action

CUTTING-EDGE CONTENT AND GUIDANCE FROM A DATA WAREHOUSING EXPERT—NOW EXPANDED TO REFLECT FIELD TRENDS Data warehousing has revolutionized the way businesses in a wide variety of industries perform analysis and make strategic decisions. Since the first edition of Data Warehousing Fundamentals, numerous enterprises have implemented data warehouse systems and reaped the benefits. Now, this new, revised edition covers the essential fundamentals of data warehousing and business intelligence as well as significant recent trends in the field. The author provides an enhanced, comprehensive overview of data warehousing together with in-depth explanations of critical issues in planning, design, deployment, and ongoing maintenance. IT professionals eager to get into the field of modern data warehousing, data warehouse architecture and infrastructure, and the various methods for information delivery. This practical Second Edition highlights the areas of data warehousing and business intelligence where high-impact technological progress has been made. Discussions on developments include data marts, real-time information delivery, data visualization, and "integrated" and "near-real-time." However, the organization's ability to successfully exploit their data is determined by the degree of strategic alignment. As such, this study poses the question: how can a data warehouse be successfully and demonstrably aligned to an organization's strategic objectives? This thesis demonstrates that strategic alignment can be achieved by following a new "top down" data warehouse implementation framework, the Configuration Approach, which is based upon determining an organization's target configuration. This was achieved by employing Miles and Snow's Ideal Types to formulate a questionnaire that reveals an organization's target configuration in terms of its approach to the Entrepreneurial, Administration and Information Systems challenges. Crucially, this thesis also provides the means to choose a data warehouse architecture that is wholly based on the organization's target configuration. The Configuration Approach was evaluated using a single case study undergoing a period of strategic transformation where the implementation of a data warehouse was key to its strategic ambitions. The case study illustrated how it is possible to articulate an organization's strategic configuration, which becomes the key driver for building a warehouse that demonstrably supports the resolution of its Entrepreneurial and Administration challenges. Significantly, the case study also provides a unique opportunity to demonstrate how the target configuration helps organizations to make the right choice of data warehouse architecture to satisfy the Information Systems challenge. In this case, the Configuration Approach provides a basis for challenging the architectural choices made by a consultancy on behalf of the participating organization. Accordingly, it can be asserted that data warehouses are strategic investments, if implemented using the Configuration Approach.

Modeling the Agile Data Warehouse with Data Vault

Data Architecture: A Primer for the Data Scientist

Data Warehousing Fundamentals for IT Professionals

Data Warehousing For Dummies

Fundamentals of Data Warehousing

Completely revised, expanded, and updated, this second edition gives extensive new coverage of data integration, management, indexing, cleansing, and transformation. The book covers powerful new multi-dimensional front-ends and conversion tools and gives detailed coverage of lifecycle issues. Unlike popular belief, Data Warehouse is not a single tool but a collection of software tools. A data warehouse will collect data from diverse sources into a single database. Using Business Intelligence tools, meaningful insights are drawn from this data. The best thing about "Learn Data Warehousing in 1 Day" is that it is small and can be completed in a day. With this e-book, you will be enough knowledge to contribute and participate in a data warehouse implementation project. The book covers upcoming and promising technologies like Data Lakes, Data Mart, ELT (Extract Load Transform) amongst others. Following are detailed topics included in the book Table of Content Chapter 1: What Is Data Warehouse? 1. What is Data Warehouse? 2. Types of Data Warehouse 3. Who needs Data warehouse? 4. Why We Need Data Warehouse? 5. Data Warehouse Tools Chapter 2: Data Warehouse Architecture 1. Characteristics of Data warehouse 2. Data Warehouse Architectures 3. Datawarehouse Components 4. Query Tools Chapter 3: ETL Process 1. What is ETL? 2. Why do you need ETL? 3. ETL Process 4. ETL tools Chapter 4: ETL Vs ETLT 1. What is ETLT? 2. Difference between ETL vs. ETLT Chapter 5: Data Modeling 1. What is Data Modelling? 2. Types of Data Models 3. Characteristics of a physical data model Chapter 6: OLAP 1. What is Online Analytical Processing? 2. Types of OLAP systems 3. Advantages and Disadvantages of OLAP Chapter 7: Multidimensional Olap (MOLAP) 1. What is MOLAP? 2. MOLAP Architecture 3. MOLAP Tools Chapter 8: OLAP Vs OLTP 1. What is the meaning of OLAP? 2. What is the meaning of OLTP? 3. Difference between OLTP and OLAP Chapter 9: Dimensional Modeling 1. What is Dimensional Model? 2. Elements of Dimensional Data Model 3. Attributes 4.

Difference between Dimension table vs. Fact table 5. Steps of Dimensional Modelling Chapter 10: Star and Snowflake Schema 1. What is Multidimensional schemas? 2. What is a Star Schema? 3. What is a Snowflake Schema? 4. Difference between Star Schema and Snowflake Chapter 11: Data Mart 1. What is Data Mart? 2. Type of Data Mart 3. Steps in Implementing a Datamart Chapter 12: Data Mart Vs Data Warehouse 1. What is Data Warehouse? 2. What is Data Warehouse? 3. Differences between a Data Warehouse and a Data Mart Chapter 13: Data Lake 1. What is Data Lake? 2. Data Lake Concepts 4. Maturity stages of Data Lake Chapter 14: Data Lake Vs Data Warehouse 1. What is Data Warehouse? 2. What is Data Lake? 3. Key Difference between the Data Lake and Data Warehouse Chapter 15: What Is Business Intelligence? 1. What Is Business Intelligence? 2. Why Is BI Important? 3. How Business Intelligence systems are Implemented? 4. Four Types of BI users Chapter 16: Data Mining 1. What is Data Mining? 2. Types of Data 3. Data Mining Process 4. Modelling 5. Data Mining Techniques Chapter 17: Data Warehousing Vs Data Mining 1. What is Data Warehouse? 2. What Is Data Mining? 3. Difference between Data mining and Data Warehousing?

At 70 terabytes and growing, Wal-Mart's data warehouse is still the world's largest, most ambitious, and arguably most successful commercial database. Written by one of the key figures in its design and construction, Data Warehousing: Using the Wal-Mart Model gives you an insider's view of this enormous project. Continuously drawing from this example, the author teaches you the general principles and specific techniques you need to understand to be a valuable part of your organization's own data warehouse project, however large or small. You'll emerge with a practical understanding of both the business and technical aspects of building a data warehouse for storing and accessing data in a strategically useful way. What's further sets this book apart is its focus on the informational needs of retail companies—including both market and organizational issues that affect the data's collection and use. If retail is your field, this book will prove especially valuable as you develop and implement your company's ideal data warehouse solution. * Written by a member of the team of four engineers who designed and built the Wal-Mart Data Warehouse database, a team whose database design was recognized internally in 1991 by Wal-Mart with the company's Team Innovational Technical award. * Provides essential information for project managers, consultants, data warehouse managers, and data architects. * Takes an in-depth Look at a wide range of technical issues, including architecture, construction approaches, tool selection, database system selection, and maintenance. * Addresses issues specific to retail business: vendors, inventory, sales analysis, geography, article categories, and more. * Explains how to determine business requirements at the outset of the project and how to develop return on investment analyses after the warehouse has been brought online.

Key takeaways from the book: Data Warehouse Architecture and Infrastructure, and the various methods for information delivery. This practical Second Edition highlights the areas of data warehousing and business intelligence where high-impact technological progress has been made. Discussions on developments include data marts, real-time information delivery, data visualization, and "integrated" and "near-real-time." However, the organization's ability to successfully exploit their data is determined by the degree of strategic alignment. As such, this study poses the question: how can a data warehouse be successfully and demonstrably aligned to an organization's strategic objectives? This thesis demonstrates that strategic alignment can be achieved by following a new "top down" data warehouse implementation framework, the Configuration Approach, which is based upon determining an organization's target configuration. This was achieved by employing Miles and Snow's Ideal Types to formulate a questionnaire that reveals an organization's target configuration in terms of its approach to the Entrepreneurial, Administration and Information Systems challenges. Crucially, this thesis also provides the means to choose a data warehouse architecture that is wholly based on the organization's target configuration. The Configuration Approach was evaluated using a single case study undergoing a period of strategic transformation where the implementation of a data warehouse was key to its strategic ambitions. The case study illustrated how it is possible to articulate an organization's strategic configuration, which becomes the key driver for building a warehouse that demonstrably supports the resolution of its Entrepreneurial and Administration challenges. Significantly, the case study also provides a unique opportunity to demonstrate how the target configuration helps organizations to make the right choice of data warehouse architecture to satisfy the Information Systems challenge. In this case, the Configuration Approach provides a basis for challenging the architectural choices made by a consultancy on behalf of the participating organization. Accordingly, it can be asserted that data warehouses are strategic investments, if implemented using the Configuration Approach.

This book presents the first comparative review of the state of the art and the best current practices of data warehouses. It covers source and data integration, multidimensional aggregation, query optimization, metadata management, quality assessment, and design optimization. A conceptual framework is presented by which the architecture and quality of a data warehouse can be assessed and improved using enriched metadata management combined with advanced techniques from databases, business modeling, and artificial intelligence.

This book presents the first comparative review of the state of the art and the best current practices of data warehouses. It covers source and data integration, multidimensional aggregation, query optimization, metadata management, quality assessment, and design optimization. A conceptual framework is presented by which the architecture and quality of a data warehouse can be assessed and improved using enriched metadata management combined with advanced techniques from databases, business modeling, and artificial intelligence.

This book presents the first comparative review of the state of the art and the best current practices of data warehouses. It covers source and data integration, multidimensional aggregation, query optimization, metadata management, quality assessment, and design optimization. A conceptual framework is presented by which the architecture and quality of a data warehouse can be assessed and improved using enriched metadata management combined with advanced techniques from databases, business modeling, and artificial intelligence.

This book presents the first comparative review of the state of the art and the best current practices of data warehouses. It covers source and data integration, multidimensional aggregation, query optimization, metadata management, quality assessment, and design optimization. A conceptual framework is presented by which the architecture and quality of a data warehouse can be assessed and improved using enriched metadata management combined with advanced techniques from databases, business modeling, and artificial intelligence.

Today, the world is trying to create and educate data scientists because of the phenomenon

Big Data fits gives the data scientist the necessary context for how pieces of the puzzle should fit together. Most references on Big Data look at only one tiny part of a much larger whole. Until data gathered can be put into an existing framework or architecture it can't be used to its full potential. Data Architecture a Primer for the Data Scientist addresses the larger architectural picture of how Big Data fits with the existing information infrastructure, an essential topic for the data scientist. Drawing upon years of practical experience and using numerous examples and an easy to understand framework. W.H. Inmon, and Daniel Linstedt define the importance of data architecture and how it can be used effectively to harness big data within existing systems. You'll be able to: Turn textual information into a form that can be analyzed by standard tools. Make the connection between analytics and Big Data Understand how Big Data fits within an existing systems environment Conduct analytics on repetitive and non-repetitive data Discusses the value in Big Data that is often overlooked, non-repetitive data, and why there is significant business value in using it Shows how to turn textual information into a form that can be analyzed by standard tools. Explains how Big Data fits within an existing systems environment Presents new opportunities that are afforded by the advent of Big Data Demystifies the murky waters of repetitive and non-repetitive data in Big Data.

PLEASE PROVIDE COURSE INFORMATION PLEASE PROVIDE

Object-oriented Data Warehouse Design

Data Architecture

Data Warehouse

The Modern Data Warehouse in Azure

DATA WAREHOUSING MOBILE CODE D

Using the Wal-Mart Model

Learn essential techniques from data warehouse legend Bill Inmon on how to build the reporting environment your business needs now! Answers for many valuable business questions hide in text. How well can your existing reporting environment extract the necessary text from email, spreadsheets, and documents, and put it in a useful format for analytics and reporting? Transforming the traditional data warehouse into an efficient unstructured data warehouse requires additional skills from the analyst, architect, designer, and developer. This book will prepare you to successfully implement an unstructured data warehouse and, through clear explanations, examples, and case studies, you will learn new techniques and tips to successfully obtain and analyze text. Master these ten objectives: • Build an unstructured data warehouse using the 11-step approach • Integrate text and describe it in terms of homogeneity, relevance, medium, volume, and structure • Overcome challenges including blather, the Tower of Babel, and lack of natural relationships • Avoid the Data Junkyard and combat the "Spider's Web" • Reuse techniques perfected in the traditional data warehouse and Data Warehouse 2.0, including iterative development • Apply essential techniques for textual Extract, Transform, and Load (ETL) such as phrase recognition, stop word filtering, and synonym replacement • Design the Document Inventory system and link unstructured text to structured data • Leverage indexes for efficient text analysis and taxonomies for useful external categorization • Manage large volumes of data using advanced techniques such as backward pointers • Evaluate technology choices suitable for unstructured data processing, such as data warehouse appliances The following outline briefly describes each chapter's content: • Chapter 1 defines unstructured data and explains why text is the main focus of this book. The sources for text, including documents, email, and spreadsheets, are described in terms of factors such as homogeneity, relevance, and structure. • Chapter 2 addresses the challenges one faces when managing unstructured data. These challenges include volume, blather, the Tower of Babel, spelling, and lack of natural relationships. Learn how to avoid a data junkyard, which occurs when unstructured data is not properly integrated into the data warehouse. This chapter emphasizes the importance of storing integrated unstructured data in a relational structure. We are cautioned on both the commonality and dangers associated with text based on paper. • Chapter 3 begins with a timeline of applications, highlighting their evolution over the decades. Eventually, powerful yet siloed applications created a "spider's web" environment. This chapter describes how data warehouses solved many problems, including the creation of corporate data, the ability to get out of the maintenance backlog conundrum, and greater data integrity and data accessibility. There were problems, however, with the data warehouse that were addressed in Data Warehouse 2.0 (DW 2.0), such as the inevitable data lifecycle. This chapter discusses the DW 2.0 architecture, which leads into the role of the unstructured data warehouse. The unstructured data warehouse is defined and benefits are given. There are several features of the conventional data warehouse that can be leveraged for the unstructured data warehouse, including ETL processing, textual integration, and iterative development. • Chapter 4 focuses on the heart of the unstructured data warehouse: Textual Extract, Transform, and Load (ETL). This chapter has separate sections on extracting text, transforming text, and loading text. The chapter emphasizes the issues around source data. There are a wide variety of sources, and each of the sources has its own set of considerations. Extracting pointers are provided, such as reading documents only once and recognizing common and different file types. Transforming text requires addressing many considerations discussed in this chapter, including phrase recognition, stop word filtering, and synonym replacement. Loading text is the final step. There are important points to understand here, too, that are explained in this chapter, such as the importance of the thematic approach and knowing how to handle large volumes of data. Two ETL examples are provided, one on email and one on spreadsheets. • Chapter 5 describes the 11 steps required to develop the unstructured data warehouse. The methodology explained in this chapter is a combination of both traditional system development lifecycle and spiral approaches. • Chapter 6 describes how to inventory documents for maximum analysis value, as well as link the unstructured text to structured data for even greater value. The Document Inventory is discussed, which is similar to a library card catalog used for organizing corporate documents. This chapter explores ways of linking unstructured text to structured data. The emphasis is on taking unstructured data and reducing it into a form of data that is structured. Related concepts to linking, such as probabilistic linkages and dynamic linkages, are discussed. • Chapter 7 goes through each of the different types of indexes necessary to make text analysis efficient. Indexes range from simple indexes, which are fast to create and are good if the analyst really knows what needs to be analyzed before the indexing process begins, to complex combined indexes, which can be made up of any and all of the other kinds of indexes. • Chapter 8 explains taxonomies and how they can be used within the unstructured data warehouse. Both simple and complicated taxonomies are discussed. Techniques to help the reader leverage taxonomies, including using preferred taxonomies, external categorization, and cluster analysis are described. Real world problems are raised, including the possibilities of encountering hierarchies, multiple types, and recursion. The chapter ends with a discussion comparing a taxonomy with a data model. • Chapter 9 explains ways of coping with large amounts of unstructured data. Techniques such as keeping the unstructured data at its source and using backward pointers are discussed. The chapter explains why iterative development is so important. Ways of reducing the amount of data are presented, including screening and removing extraneous data, as well as parallelizing the workload. • Chapter 10 focuses on challenges and some technology choices that are suitable for unstructured data processing. The traditional data warehouse processing technology is reviewed. In addition, the data warehouse appliance is discussed. • Chapters 11, 12, and 13 put all of the previously discussed techniques and approaches in context through three case studies: the Ablatz Medical Group, the Eastern Hills Oil Company, and the Amber Oil Company. Data Modeling for Agile Data Warehouse using Data Vault Modeling Approach. Includes Enterprise Data Warehouse Architecture. This is a complete guide to the data vault data modeling approach. The book also includes business and program considerations for the agile data warehousing and business intelligence program. There are over 200 diagrams and figures concerning modeling, core business concepts, architecture, business alignment, semantics, and modeling comparisons with 3NF and Dimensional modeling.

Data Warehouse From Architecture to Implementation Addison-Wesley Professional

Techniques for building modern cloud data warehousing solutions

The Complete Guide to Dimensional Modeling