

## The Data Model Toolkit Simple Skills To Model The Real World Data Architecture Fundamentals

Are you looking to make better use of data captured within your organisation or want to learn more about how Data Architecture can transform your operations? Answering these questions is at the very heart of Navigating the Data Architecture Landscape. By reading this book you will learn how to: Introduce or improve the Data Architecture function of your organisation Enhance your skills in this domain to deliver more from your data. You may be wondering how a book can do this if it knows nothing about where you are now, or where you want to be? It can, because by leveraging its principles you will discover how to create optimised potential routes to achieve your own Data Architectural objectives. Basic building blocks, concepts and models are defined, enabling you to create new or adapt existing frameworks appropriate for any data landscape. Practical tips and suggestions are also detailed throughout, helping you gain immediate improvements from the way you work and enhance the benefits your organisation can derive from its data. So if you are a Data Architect or deal with data in your organisation and want to learn how to transform the positive yield from its data, then this book is a must read for you!

“David has been there and dealt with the issues, which is why this book is an outstanding resource for Data Architects and indeed anyone dealing with the serious challenges of an enterprise data landscape.” – Richard Rendell, Technical Services Director, AgeSmart “An essential read for anyone wishing to practically achieve more benefit from data for their organisation within today’s constraints.” – Reem Zahran - Director, Offering Development, IMS Health “This book provides a comprehensive set of tools enabling you to improve the business outcomes from your organisation’s use of data.” – Andrew Rowland, Global Head Database Engineering, UBS This book is an essential read for Data Architects or indeed anyone wanting to improve the benefit that their organisation can derive from its data usage. It does this by providing principles and models that are appropriate to use within any framework, or even the absence of one. The book is designed to be practical and contains many tips and suggestions as well as examples that can be used as the basis for the reader's own Data Architectural definitions. The breadth of the book covers contemporary themes for Data Architecture and the chapters include; Data Modelling, Enterprise Data Models, Data Governance, Master Data Management and Big Data

Data Resource Data provides the complete detailed data resource model for understanding and managing data as a critical resource of the organization.

". . . one of the definitive books of our industry. If you take the time to read only one professional book, make it this book." -W. H. Inmon One of the most dramatic new developments in database design, the dimensional data warehouse is a powerful database model that significantly enhances managers' ability to quickly analyze large, multidimensional data sets. Written by the leading proponent of this revolutionary new approach, this valuable book/CD toolkit outfits you with all the nuts-and-bolts information you need to design, build, manage, and use dimensional data warehouses for virtually any type of business application, as well as software for querying dimensional data warehouses. Employing many real-life case studies of data warehouses, Ralph Kimball provides clear-cut guidelines on how to model data and design data warehouses to support advanced multidimensional decision support systems. Beginning with the relatively simple example of a data warehouse for a grocery store, he progresses, step-by-step, through an increasingly complex array of business applications in retail, manufacturing, banking, insurance, subscriptions, and airline reservations. By the end of the book, you will have mastered the full range of powerful techniques for creating, controlling, and navigating dimensional business databases that are easy to understand and navigate. On the CD-ROM you'll find: \* Software for querying dimensional data warehouses. \* Working models of all the databases described in the book.

This book offers a highly accessible introduction to natural language processing, the field that supports a variety of language technologies, from predictive text and email filtering to automatic summarization and translation. With it, you'll learn how to write Python programs that work with large collections of unstructured text. You'll access richly annotated datasets using a comprehensive range of linguistic data structures, and you'll understand the main algorithms for analyzing the content and structure of written communication. Packed with examples and exercises, Natural Language Processing with Python will help you: Extract information from unstructured text, either to guess the topic or identify "named entities" Analyze linguistic structure in text, including parsing and semantic analysis Access popular linguistic databases, including WordNet and treebanks Integrate techniques drawn from fields as diverse as linguistics and artificial intelligence This book will help you gain practical skills in natural language processing using the Python programming language and the Natural Language Toolkit (NLTK) open source library. If you're interested in developing web applications, analyzing multilingual news sources, or documenting endangered languages -- or if you're simply curious to have a programmer's perspective on how human language works -- you'll find Natural Language Processing with Python both fascinating and immensely useful.

What is this book about? Professional Java builds upon Ivor Horton's Beginning Java to provide the reader with an understanding of how professionals use Java to develop software solutions. Pro Java starts with an overview of best methods and tools for developing Java applications. It then examines the the more sophisticated and nuanced parts of the Java JDK. The final and most extensive part of the book shows how to implement these ideas to build real-world applications, using both Java APIs as well as related Java open source tools. In short, this book provides a comprehensive treatment of the professional Java development process, without losing focus in exhaustive coverage of isolated features and APIs.

Did you ever try getting Businesspeople and IT to agree on the project scope for a new application? Or try getting Marketing and Sales to agree on the target audience? Or try bringing new team members up to speed on the hundreds of tables in your data warehouse — without them dozing off? Whether you are a businessperson or an IT professional,

you can be the hero in each of these and hundreds of other scenarios by building a High-Level Data Model. The High-Level Data Model is a simplified view of our complex environment. It can be a powerful communication tool of the key concepts within our application development projects, business intelligence and master data management programs, and all enterprise and industry initiatives. Learn about the High-Level Data Model and master the techniques for building one, including a comprehensive ten-step approach and hands-on exercises to help you practice topics on your own. In this book, we review data modeling basics and explain why the core concepts stored in a high-level data model can have significant business impact on an organization. We explain the technical notation used for a data model and walk through some simple examples of building a high-level data model. We also describe how data models relate to other key initiatives you may have heard of or may be implementing in your organization. This book contains best practices for implementing a high-level data model, along with some easy-to-use templates and guidelines for a step-by-step approach. Each step will be illustrated using many examples based on actual projects we have worked on. Names have been changed to protect the innocent, but the pain points and lessons have been preserved. One example spans an entire chapter and will allow you to practice building a high-level data model from beginning to end, and then compare your results to ours. Building a high-level data model following the ten step approach you'll read about is a great way to ensure you will retain the new skills you learn in this book. As is the case in many disciplines, using the right tool for the right job is critical to the overall success of your high-level data model implementation. To help you in your tool selection process, there are several chapters dedicated to discussing what to look for in a high-level data modeling tool and a framework for choosing a data modeling tool, in general. This book concludes with a real-world case study that shows how an international energy company successfully used a high-level data model to streamline their information management practices and increase communication throughout the organization—between both businesspeople and IT. Data modeling is one of the under-exploited, and potentially very valuable, business capabilities that are often hidden away in an organization's Information Technology department. Data Modeling for the Business highlights both the resulting damage to business value, and the opportunities to make things better. As an easy-to follow and comprehensive guide on the 'why' and 'how' of data modeling, it also reminds us that a successful strategy for exploiting IT depends at least as much on the information as the technology. Chris Potts, Corporate IT Strategist and Author of *frulTion: Creating the Ultimate Corporate Strategy for Information Technology* One of the most critical systems issues is aligning business with IT and fulfilling business needs using data models. The authors of *Data Modeling for the Business* do a masterful job at simply and clearly describing the art of using data models to communicate with business representatives and meet business needs. The book provides many valuable tools, analogies, and step-by-step methods for effective data modeling and is an important contribution in bridging the much needed connection between data modeling and realizing business requirements. Len Silverston, author of *The Data Model Resource Book* series

For many researchers, Python is a first-class tool mainly because of its libraries for storing, manipulating, and gaining insight from data. Several resources exist for individual pieces of this data science stack, but only with the *Python Data Science Handbook* do you get them all—IPython, NumPy, Pandas, Matplotlib, Scikit-Learn, and other related tools. Working scientists and data crunchers familiar with reading and writing Python code will find this comprehensive desk reference ideal for tackling day-to-day issues: manipulating, transforming, and cleaning data; visualizing different types of data; and using data to build statistical or machine learning models. Quite simply, this is the must-have reference for scientific computing in Python. With this handbook, you'll learn how to use: IPython and Jupyter: provide computational environments for data scientists using Python NumPy: includes the ndarray for efficient storage and manipulation of dense data arrays in Python Pandas: features the DataFrame for efficient storage and manipulation of labeled/columnar data in Python Matplotlib: includes capabilities for a flexible range of data visualizations in Python Scikit-Learn: for efficient and clean Python implementations of the most important and established machine learning algorithms

*Data Modeling Essentials, Third Edition*, covers the basics of data modeling while focusing on developing a facility in techniques, rather than a simple familiarization with "the rules". In order to enable students to apply the basics of data modeling to real models, the book addresses the realities of developing systems in real-world situations by assessing the merits of a variety of possible solutions as well as using language and diagramming methods that represent industry practice. This revised edition has been given significantly expanded coverage and reorganized for greater reader comprehension even as it retains its distinctive hallmarks of readability and usefulness. Beginning with the basics, the book provides a thorough grounding in theory before guiding the reader through the various stages of applied data modeling and database design. Later chapters address advanced subjects, including business rules, data warehousing, enterprise-wide modeling and data management. It includes an entirely new section discussing the development of logical and physical modeling, along with new material describing a powerful technique for model verification. It also provides an excellent resource for additional lectures and exercises. This text is the ideal reference for data modelers, data architects, database designers, DBAs, and systems analysts, as well as undergraduate and graduate-level students looking for a real-world perspective. Thorough coverage of the fundamentals and relevant theory. Recognition and support for the creative side of the process. Expanded coverage of applied data modeling includes new chapters on logical and physical database design. New material describing a powerful technique for model verification. Unique coverage of the practical and human aspects of modeling, such as working with business specialists, managing change, and resolving conflict.

This is the digital version of the printed book (Copyright © 1996). Learning the basics of a modeling technique is not the same as learning how to use and apply it. To develop a data model of an organization is to gain insights into its nature that do not come easily. Indeed, analysts are often expected to understand subtleties of an organization's structure that may have evaded people who have worked there for years. Here's help for those analysts who have learned the basics of data modeling (or "entity/relationship modeling") but who need to obtain the insights

required to prepare a good model of a real business. Structures common to many types of business are analyzed in areas such as accounting, material requirements planning, process manufacturing, contracts, laboratories, and documents. In each chapter, high-level data models are drawn from the following business areas: The Enterprise and Its World The Things of the Enterprise Procedures and Activities Contracts Accounting The Laboratory Material Requirements Planning Process Manufacturing Documents Lower-Level Conventions The Data Model Toolkit Simple Skills To Model The Real World Paragon Publishing

Best practices and invaluable advice from world-renowned data warehouse experts In this book, leading data warehouse experts from the Kimball Group share best practices for using the upcoming "Business Intelligence release" of SQL Server, referred to as SQL Server 2008 R2. In this new edition, the authors explain how SQL Server 2008 R2 provides a collection of powerful new tools that extend the power of its BI toolset to Excel and SharePoint users and they show how to use SQL Server to build a successful data warehouse that supports the business intelligence requirements that are common to most organizations. Covering the complete suite of data warehousing and BI tools that are part of SQL Server 2008 R2, as well as Microsoft Office, the authors walk you through a full project lifecycle, including design, development, deployment and maintenance. Features more than 50 percent new and revised material that covers the rich new feature set of the SQL Server 2008 R2 release, as well as the Office 2010 release Includes brand new content that focuses on PowerPivot for Excel and SharePoint, Master Data Services, and discusses updated capabilities of SQL Server Analysis, Integration, and Reporting Services Shares detailed case examples that clearly illustrate how to best apply the techniques described in the book The accompanying Web site contains all code samples as well as the sample database used throughout the case studies The Microsoft Data Warehouse Toolkit, Second Edition provides you with the knowledge of how and when to use BI tools such as Analysis Services and Integration Services to accomplish your most essential data warehousing tasks.

From three design partners at Google Ventures, a unique five-day process--called the sprint--for solving tough problems using design, prototyping, and testing ideas with customers. Database Modeling and Design, Fifth Edition, focuses on techniques for database design in relational database systems. This extensively revised fifth edition features clear explanations, lots of terrific examples and an illustrative case, and practical advice, with design rules that are applicable to any SQL-based system. The common examples are based on real-life experiences and have been thoroughly class-tested. This book is immediately useful to anyone tasked with the creation of data models for the integration of large-scale enterprise data. It is ideal for a stand-alone data management course focused on logical database design, or a supplement to an introductory text for introductory database management. In-depth detail and plenty of real-world, practical examples throughout Loaded with design rules and illustrative case studies that are applicable to any SQL, UML, or XML-based system Immediately useful to anyone tasked with the creation of data models for the integration of large-scale enterprise data.

This book gives a general picture of research-driven activities related to location and map-based services. The interdisciplinary character of the topic leads to a variety of contributions with backgrounds from academia to business and from computer science to geodesy. While cartography is aiming at efficient communication of spatial information, the development and availability of technologies like mobile networking, mobile devices or short-range sensors lead to interesting new possibilities of achieving this aim. By trying to make use of the available technologies, a variety of related disciplines looks specifically at user-centered and context-aware system development, especially in wayfinding and navigation systems.

This book is for all data modelers, data architects, and database designers?be they novices who want to learn what's involved in data modeling, or experienced modelers who want to brush up their skills. A novice will not only gain an overview of data modeling, they will also learn how to follow the data modeling process, including the activities required for each step. The experienced practitioner will discover (or rediscover) techniques to ensure that data models accurately reflect business requirements. This book describes rigorous yet easily implemented approaches to:

- modeling of business information requirements for review by business stakeholders before development of the logical data model
- normalizing data, based on simple questions rather than the formal definitions which many modelers find intimidating
- naming and defining concepts and attributes
- modeling of time-variant data
- documenting business rules governing both the real world and data
- data modeling in an Agile project
- managing data model change in any type of project
- transforming a business information model to a logical data model against which developers can code
- implementing the logical data model in a traditional relational DBMS, an SQL:2003-compliant DBMS, an object-relational DBMS, or in XML.

Part 1 describes business information models in-depth, including:

- the importance of modeling business information requirements before embarking on a logical data model
- business concepts (entity classes)
- attributes of business concepts
- attribute classes as an alternative to DBMS data types
- relationships between business concepts
- time-variant data
- generalization and specialization of business concepts
- naming and defining the components of the business information model
- business rules governing data, including a distinction between real-world rules and data rules.

Part 2 journeys from requirements to a working data resource, covering:

- sourcing data requirements
- developing the business information model
- communicating it to business stakeholders for review, both as diagrams and verbally
- managing data model change
- transforming the business information model into a logical data model of stored data for implementation in a relational or object-relational DBMS
- attribute value representation and data constraints (important but often overlooked)
- modeling data vault, dimensional and XML data.

As data management and integration continue to evolve rapidly, storing all your data in one place, such as a data warehouse, is no longer scalable. In the very near future, data will need to be distributed and available for several technological solutions. With this practical book, you'll learn how to migrate your enterprise from a complex and tightly coupled data landscape to a more flexible architecture ready for the modern world of data consumption. Executives, data architects, analytics teams, and compliance and governance staff will learn how to build a modern scalable data landscape using the Scaled Architecture, which you can introduce incrementally without a large upfront investment. Author Piethein Strengholt provides blueprints, principles, observations, best practices, and patterns to get you up to speed. Examine data management trends, including technological developments, regulatory requirements, and privacy concerns Go deep into the Scaled Architecture and learn how the pieces fit together Explore data governance and data security, master data management, self-service data marketplaces, and the importance of metadata

Behavioural insights (BI) are lessons derived from the behavioural and social sciences, including decision making, psychology, cognitive science, neuroscience, organisational

and group behaviour.

The subject of hypertext and hypermedia has been witnessing a dramatic rise in interest over the last three years both from the academic and the industrial communities. This volume presents the proceedings of the European Conference on Hypertext (ECHT90) conference held in Paris in November 1990, where researchers, developers, and users were able to meet and discuss the theme of hypertext and hypermedia. This volume will be of interest to a broad spectrum of professionals ranging from pure theoreticians of hypergraphs and graph grammars via system developers for electronic publishing to end-users of hypertext applications such as medical information systems and computer aided design.

Data Modeling Made Simple with CA ERwin Data Modeler r8 will provide the business or IT professional with a practical working knowledge of data modeling concepts and best practices, and how to apply these principles with CA ERwin Data Modeler r8. You'll build many CA ERwin data models along the way, mastering first the fundamentals and later in the book the more advanced features of CA ERwin Data Modeler. This book combines real-world experience and best practices with down to earth advice, humor, and even cartoons to help you master the following ten objectives: 1. Understand the basics of data modeling and relational theory, and how to apply these skills using CA ERwin Data Modeler 2. Read a data model of any size and complexity with the same confidence as reading a book 3. Understand the difference between conceptual, logical, and physical models, and how to effectively build these models using CA ERwin's Data Modelers Design Layer Architecture 4. Apply techniques to turn a logical data model into an efficient physical design and vice-versa through forward and reverse engineering, for both 'top down' and bottom-up design 5. Learn how to create reusable domains, naming standards, UDPs, and model templates in CA ERwin Data Modeler to reduce modeling time, improve data quality, and increase enterprise consistency 6. Share data model information with various audiences using model formatting and layout techniques, reporting, and metadata exchange 7. Use the new workspace customization features in CA ERwin Data Modeler r8 to create a workflow suited to your own individual needs 8. Leverage the new Bulk Editing features in CA ERwin Data Modeler r8 for mass metadata updates, as well as import/export with Microsoft Excel 9. Compare and merge model changes using CA ERwin Data Modelers Complete Compare features 10. Optimize the organization and layout of your data models through the use of Subject Areas, Diagrams, Display Themes, and more Section I provides an overview of data modeling: what it is, and why it is needed. The basic features of CA ERwin Data Modeler are introduced with a simple, easy-to-follow example. Section II introduces the basic building blocks of a data model, including entities, relationships, keys, and more. How-to examples using CA ERwin Data Modeler are provided for each of these building blocks, as well as 'real world' scenarios for context. Section III covers the creation of reusable standards, and their importance in the organization. From standard data modeling constructs such as domains to CA ERwin-specific features such as UDPs, this section covers step-by-step examples of how to create these standards in CA ERwin Data Modeling, from creation, to template building, to sharing standards with end users through reporting and queries. Section IV discusses conceptual, logical, and physical data models, and provides a comprehensive case study using CA ERwin Data Modeler to show the interrelationships between these models using CA ERwin's Design Layer Architecture. Real world examples are provided from requirements gathering, to working with business sponsors, to the hands-on nitty-gritty details of building conceptual, logical, and physical data models with CA ERwin Data Modeler r8. From the Foreword by Tom Bilcze, President, CA Technologies Modeling Global User Community: Data Modeling Made Simple with CA ERwin Data Modeler r8 is an excellent resource for the ERwin community. The data modeling community is a diverse collection of data professionals with many perspectives of data modeling and different levels of skill and experience. Steve Hoberman and Donna Burbank guide newbie modelers through the basics of data modeling and CA ERwin r8. Through the liberal use of illustrations, the inexperienced data modeler is graphically walked through the components of data models and how to create them in CA ERwin r8. As an experienced data modeler, Steve and Donna give me a handbook for effectively using the new and enhanced features of this release to bring my art form to life. The book delves into advanced modeling topics and techniques by continuing the liberal use of illustrations. It speaks to the importance of a defined data modeling architecture with soundly modeled data to assist the enterprise in understanding of the value of data. It guides me in applying the finishing touches to my data designs.

Data Modeling Made Simple with PowerDesigner will provide the business or IT professional with a practical working knowledge of data modeling concepts and best practices, and how to apply these principles with PowerDesigner. You'll build many PowerDesigner data models along the way, increasing your skills first with the fundamentals and later with more advanced feature of PowerDesigner. This book combines real-world experience and best practices to help you master the following ten objectives: This book has ten key objectives for you, the reader: 1. You will know when a data model is needed and which PowerDesigner models are the most appropriate for each situation 2. You will be able to read a data model of any size and complexity with the same confidence as reading a book 3. You will know when to apply and how to make use of all the key features of PowerDesigner 4. You will be able to build, step-by-step in PowerDesigner, a pyramid of linked data models, including a conceptual data model, a fully normalized relational data model, a physical data model, and an easily navigable dimensional model 5. You will be able to apply techniques such as indexing, transforms, and forward engineering to turn a logical data model into an efficient physical design 6. You will improve data governance and modeling consistency within your organization by leveraging features such as PowerDesigner's reference models, Glossary, domains, and model comparison and model mapping techniques 7. You will know how to utilize dependencies and traceability links to assess the impact of change 8. You will know how to integrate your PowerDesigner models with externally-managed files, including the import and export of data using Excel and Requirements documents 9. You will know where you can take advantage of the entire PowerDesigner model set, to increase the success rate of corporate-wide

initiatives such as business intelligence and enterprise resource planning (ERP) 10. You will understand the key differentiators between PowerDesigner and other data modeling tools you may have used before This book contains seven sections: Section I introduces data modeling, along with its purpose and variations. Section II explains all of the components on a data model including entities, data elements, relationships, and keys. Also included is a discussion of the importance of quality names and definitions for your objects. Section III explains the important role of data modeling tools, the key features required of any data modeling tool, and an introduction to the essential features of PowerDesigner. It also describes how to create and manage data modeling objects in PowerDesigner. Section IV introduces the Data Model Pyramid, then dives into the relational and dimensional subject areas, logical, and physical data models, and describes how PowerDesigner supports these models and the connections between them. Section V guides you through the creation of your own Data Model Pyramid. Section VI focuses on additional PowerDesigner features (some of which have already been introduced) that make life easier for data modelers. Learn how to get information into and out of PowerDesigner, and improve the quality of your data models with a cross-reference of key PowerDesigner features with the Data Model Scorecard®. Section VII discusses PowerDesigner topics beyond data modeling, including the XML physical model and the other types of model available in PowerDesigner.

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.

Agile Data Warehouse Design is a step-by-step guide for capturing data warehousing/business intelligence (DW/BI) requirements and turning them into high performance dimensional models in the most direct way: by modelstorming (data modeling ] brainstorming) with BI stakeholders. This book describes BEAM, an agile approach to dimensional modeling, for improving communication between data warehouse designers, BI stakeholders and the whole DW/BI development team. BEAM provides tools and techniques that will encourage DW/BI designers and developers to move away from their keyboards and entity relationship based tools and model interactively with their colleagues. The result is everyone thinks dimensionally from the outset! Developers understand how to efficiently implement dimensional modeling solutions. Business stakeholders feel ownership of the data warehouse they have created, and can already imagine how they will use it to answer their business questions. Within this book, you will learn: Agile dimensional modeling using Business Event Analysis & Modeling (BEAM ) Modelstorming: data modeling that is quicker, more inclusive, more productive, and frankly more fun! Telling dimensional data stories using the 7Ws (who, what, when, where, how many, why and how) Modeling by example not abstraction; using data story themes, not crow's feet, to describe detail Storyboarding the data warehouse to discover conformed dimensions and plan iterative development Visual modeling: sketching timelines, charts and grids to model complex process measurement - simply Agile design documentation: enhancing star schemas with BEAM dimensional shorthand notation Solving difficult DW/BI performance and usability problems with proven dimensional design patterns LawrenceCorr is a data warehouse designer and educator. As Principal of DecisionOne Consulting, he helps clients to review and simplify their data warehouse designs, and advises vendors on visual data modeling techniques. He regularly teaches agile dimensional modeling courses worldwide and has taught dimensional DW/BI skills to thousands of students. Jim Stagnitto is a data warehouse and master data management architect specializing in the healthcare, financial services, and information service industries. He is the founder of the data warehousing and data mining consulting firm Llumino.

Build a working knowledge of data modeling concepts and best practices, along with how to apply these principles with ER/Studio. This second edition includes numerous updates and new sections including an overview of ER/Studio's support for agile development, as well as a description of some of ER/Studio's newer features for NoSQL, such as MongoDB's containment structure.

Updated new edition of Ralph Kimball's groundbreaking book on dimensional modeling for data warehousing and business intelligence! The first edition of Ralph Kimball's The Data Warehouse Toolkit introduced the industry to dimensional modeling, and now his books are considered the most authoritative guides in this space. This new third edition is a complete library of updated dimensional modeling techniques, the most comprehensive collection ever. It covers new and enhanced star schema dimensional modeling patterns, adds two new chapters on ETL techniques, includes new and expanded business matrices for 12 case studies, and more. Authored by Ralph Kimball and Margy Ross, known worldwide as educators, consultants, and influential thought leaders in data warehousing and business intelligence Begins with fundamental design recommendations and progresses through increasingly complex scenarios Presents unique modeling techniques for business applications such as inventory management, procurement, invoicing, accounting, customer relationship management, big data analytics, and more Draws real-world case studies from a variety of industries, including retail sales, financial services, telecommunications, education, health care, insurance, e-commerce, and more Design dimensional databases that are easy to understand and provide fast query response with

The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, 3rd Edition.

Wouldn't it be great to understand all the data in your organisation? Just imagine being able to define, agree and manage information concepts that impact on business strategy? Then imagine that these information concepts can be linked to the physical database attributes that ultimately are used to create them. That's what this book is about. It focuses on the data model as the foundation for achieving this understanding. This book provides a framework for the enterprise data model, the business reasons behind it and the differences between conceptual, logical and physical data models. The question of how, and why, to use a data model artifact as part of the data governance toolkit for the whole enterprise is also addressed. This publication is not an in-depth manual on how to model data for a new database system or your next design project. It instead focuses at a level above these implementation projects and addresses the issues that organisations typical struggling with such as: \* How do we provide a framework within which we can manage our data assets? \* How do we develop applications that adhere to a set of data standards; without creating a nightmare of administration and governance that is both unwieldy and unusable? \* How can we get business value from our enterprise data? Chapter headings are: \* Chapter 1 - Introduction \* Chapter 2 - Information and Data \* Chapter 3 - Pillars of Value \* Chapter 4 - An Overview of Data Modelling \* Chapter 5 - Data Architecture \* Chapter 6 - The Enterprise Data Model \* Chapter 7 - Build the Model one Project at a Time \* Chapter 8 - Master Data \* Chapter 9 - Data Governance \* Chapter 10 - The Enterprise Data Framework This 2nd edition revises the original text to add extra details around key areas such as the enterprise data model framework and the pillars of value. It also improves the quality of the original text.

Proven data models that save companies' time and money developing data architectures, databases and data warehouses. The Data Model Resource Book provides a common set of data models for specific functions common to most businesses, such as sales, marketing, order processing, contracts, shipments, invoicing, work efforts, budgeting, and accounting. Readers can apply more than one data model to their own company to meet specific data needs. A CD-ROM, sold separately, contains the SQL for all of the data models contained in the book in Oracle, Sybase, Informix, Microsoft SQL Server and ANSI compatible formats.

This third volume of the best-selling "Data Model Resource Book" series revolutionizes the data modeling discipline by answering the question "How can you save significant time while improving the quality of any type of data modeling effort?" In contrast to the first two volumes, this new volume focuses on the fundamental, underlying patterns that affect over 50 percent of most data modeling efforts. These patterns can be used to considerably reduce modeling time and cost, to jump-start data modeling efforts, as standards and guidelines to increase data model consistency and quality, and as an objective source against which an enterprise can evaluate data models. Praise for The Data Model Resource Book, Volume 3 "Len and Paul look beneath the superficial issues of data modeling and have produced a work that is a must for every serious designer and manager of an IT project." —Bill Inmon, World-renowned expert, speaker, and author on data warehousing and widely recognized as the "father of data warehousing" "The Data Model Resource Book, Volume 3: Universal Patterns for Data Modeling is a great source for reusable patterns you can use to save a tremendous amount of time, effort, and cost on any data modeling effort. Len Silverston and Paul Agnew have provided an indispensable reference of very high-quality patterns for the most foundational types of data model structures. This book represents a revolutionary leap in moving the data modeling profession forward." —Ron Powell, Cofounder and Editorial Director of the Business Intelligence Network "After we model a Customer, Product, or Order, there is still more about each of these that remains to be captured, such as roles they play, classifications in which they belong, or states in which they change. The Data Model Resource Book, Volume 3: Universal Patterns for Data Modeling clearly illustrates these common structures. Len Silverston and Paul Agnew have created a valuable addition to our field, allowing us to improve the consistency and quality of our models by leveraging the many common structures within this text." —Steve Hoberman, Best-Selling Author of Data Modeling Made Simple "The large national health insurance company I work at has actively used these data patterns and the (Universal Data Models) UDM, ahead of this book, through Len Silverston's UDM Jump Start engagement. The patterns have found their way into the core of our Enterprise Information Model, our data warehouse designs, and progressively into key business function databases. We are getting to reuse the patterns across projects and are reaping benefits in understanding, flexibility, and time-to-market. Thanks so much." —David Chasteen, Enterprise Information Architect "Reusing proven data modeling design patterns means exactly that. Data models become stable, but remain very flexible to accommodate changes. We have had the fortune of having Len and Paul share the patterns that are described in this book via our engagements with Universal Data Models, LLC. These data modeling design patterns have helped us to focus on the essential business issues because we have leveraged these reusable building blocks for many of the standard design problems. These design patterns have also helped us to evaluate the quality of data models for their intended purpose. Many times there are a lot of enhancements required. Too often the very specialized business-oriented data model is also implemented physically. This may have significant drawbacks to flexibility. I'm looking forward to increasing the data modeling design pattern competence within Nokia with the help of this book." —Teemu Mattelmaki, Chief Information Architect, Nokia "Once again, Len Silverston, this time together with Paul Agnew, has made a valuable contribution to the body of knowledge about data models, and the act of building sound data models. As a professional d

A quick and reliable way to build proven databases for core business functions Industry experts raved about The Data Model Resource Book when it was first published in March 1997 because it provided a simple, cost-effective way to design databases for core business functions. Len Silverston has now revised and updated the hugely successful 1st Edition, while adding a companion volume to take care of more specific requirements of different businesses. This updated volume provides a common set of data models for specific core functions shared by most businesses like human resources management, accounting, and project management. These models are standardized and are easily

replicated by developers looking for ways to make corporate database development more efficient and cost effective. This guide is the perfect complement to The Data Model Resource CD-ROM, which is sold separately and provides the powerful design templates discussed in the book in a ready-to-use electronic format. A free demonstration CD-ROM is available with each copy of the print book to allow you to try before you buy the full CD-ROM.

Adopting the latest technological and data related innovations has caused many organisations to realise they don't have a firm grasp on their basic operational data. This is a problem that Logical Data Models are uniquely qualified to help them solve. The realisation of the need to define a Logical Data Model may be driven by any number of reasons including; trying to link Big Data Analytics to operational data, plunging into Digital Marketing, choosing the best SaaS solution, carrying out a core Data Migration, developing a Data Warehouse, enhancing Data Governance processes, or even just trying to get everyone to agree on their Product specifications! This book will provide you with the skills required to start to answer these and many similar types of questions. It is not written with a focus on IT development, so you don't need a technical background to get the most from it. But for any professional working in an organisation's data landscape, this book will provide the skills they need to define high quality and beneficial data models quickly and easily. It does this using a wealth of practical examples, tips and techniques, as well as providing checklists and templates. It is structured into three parts: The Foundations: What are the solid foundations necessary for building effective data models? The Tools: What Tools are required to enable you to specify clear, precise and accurate data model definitions? The Deliverables: What processes will you need to successfully define the models, what will they deliver, and how can we make them beneficial to the organisation? "In this data-rich era, it is even more critical for organisations to answer the question of what their data means and the value it can bring. Those who can, will gain a competitive advantage through their use of data to streamline their operations and energise their strategies. Core to revealing this meaning, is the data model that is now, more than ever, the lynchpin of success. The Data Model Toolkit provides the essential knowledge and skills that will ensure this success." – Reem Zahran, Global IT Platform Director, TNS "We work with many enterprise customers to help them transform their technology and it always starts with data. The key is a clear definition of their data quality, completeness and governance. This book shows you step by step how to define and use Data Models as powerful tools to define an organisation's data and maximise its business benefit." – John Casserly, CEO, Xceed Group

Cowritten by Ralph Kimball, the world's leading data warehousing authority, whose previous books have sold more than 150,000 copies Delivers real-world solutions for the most time- and labor-intensive portion of data warehousing-data staging, or the extract, transform, load (ETL) process Delineates best practices for extracting data from scattered sources, removing redundant and inaccurate data, transforming the remaining data into correctly formatted data structures, and then loading the end product into the data warehouse Offers proven time-saving ETL techniques, comprehensive guidance on building dimensional structures, and crucial advice on ensuring data quality

Ralph Kimball invented a data warehousing technique called 'dimensional modelling' and popularised it in his first Wiley bestseller The Data Warehouse Toolkit. Since then dimensional modelling has become the most widely accepted technique for data warehouse design. Since the first edition, Kimball has improved on his earlier techniques and created many new ones. In this second edition, he provides a comprehensive collection of all of them, from basic to advanced, and strategies for optimising data warehouse design for common business applications. He includes examples for retail sales, inventory management, procurement, orders and invoices, customer relationship management, accounting, financial services, telecommunication and utilities, health care, insurance and more. He also presents unique modelling techniques for e-commerce and shows strategies for optimising performance. A companion Web site provides updates on dimensional modelling techniques, links to related sites and source code where appropriate. This book constitutes the refereed proceedings of the 6th International Conference on Intelligent Data Analysis, IDA 2005, held in Madrid, Spain in September 2005. The 46 revised papers presented together with two tutorials and two invited talks were carefully reviewed and selected from 184 submissions. All current aspects of this interdisciplinary field are addressed; the areas covered include statistics, machine learning, data mining, classification and pattern recognition, clustering, applications, modeling, and interactive dynamic data visualization.

Data Modeling Made Simple will provide the business or IT professional with a practical working knowledge of data modeling concepts and best practices. This book is written in a conversational style that encourages you to read it from start to finish and master these ten objectives: Know when a data model is needed and which type of data model is most effective for each situation Read a data model of any size and complexity with the same confidence as reading a book Build a fully normalized relational data model, as well as an easily navigatable dimensional model Apply techniques to turn a logical data model into an efficient physical design Leverage several templates to make requirements gathering more efficient and accurate Explain all ten categories of the Data Model Scorecard Learn strategies to improve your working relationships with others Appreciate the impact unstructured data has, and will have, on our data modeling deliverables Learn basic UML concepts Put data modeling in context with XML, metadata, and agile development Book Review by Johnny Gay In this book review, I address each section in the book and provide what I found most valuable as a data modeler. I compare, as I go, how the book's structure eases the new data modeler into the subject much like an instructor might ease a beginning swimmer into the pool. This book begins like a Dan Brown novel. It even starts out with the protagonist, our favorite data modeler, lost on a dark road somewhere in France. In this case, what saves him isn't a cipher, but of all things, something that's very much like a data model in the form of a map! The author deems they are both way-finding tools. The chapters in the book are divided into 5 sections. The chapters in each section end with an exercise and a list of the key points covered to reinforce what you've learned. I find myself comparing the teaching structure of the book to the way most of us learn to swim.

The Multi-dimensional Data Modeling Toolkit represents over 15 years of hands-on experience developing multidimensional analytic applications for over a dozen companies in a variety of application areas. Written in a tutorial style, this book gives, in plain English, a step-by-step development of the defining principles of OLAP analysis through the lens of the programming language at the heart of Oracle's OLAP database option. You will find this book packed with examples, tricks and techniques, concrete illustrations of the programming elements needed to implement. The basics will all be there as well as advanced

techniques that you can use to address the most demanding requirements. OLAP will be addressed as an analysis platform. You will learn how to make business intelligence applications smarter by upping the analytical octane. You will learn both the classic applications of OLAP analysis as well as more exotic approaches. You will learn where OLAP fits in among other analytical approaches such as statistics and data mining. So whether you are a developer wanting to learn Oracle's counterpart to Microsoft's MDX, or an analyst wanting to understand the quantitative possibilities of OLAP, The Multi-dimensional Data Modeling Toolkit will show you what you need to know to go from beginner to expert in the application of OLAP analytics with Oracle OLAP DML.

While many companies ponder implementation details such as distributed processing engines and algorithms for data analysis, this practical book takes a much wider view of big data development, starting with initial planning and moving diligently toward execution. Authors Ted Malaska and Jonathan Seidman guide you through the major components necessary to start, architect, and develop successful big data projects. Everyone from CIOs and COOs to lead architects and developers will explore a variety of big data architectures and applications, from massive data pipelines to web-scale applications. Each chapter addresses a piece of the software development life cycle and identifies patterns to maximize long-term success throughout the life of your project. Start the planning process by considering the key data project types Use guidelines to evaluate and select data management solutions Reduce risk related to technology, your team, and vague requirements Explore system interface design using APIs, REST, and pub/sub systems Choose the right distributed storage system for your big data system Plan and implement metadata collections for your data architecture Use data pipelines to ensure data integrity from source to final storage Evaluate the attributes of various engines for processing the data you collect

This book will show you how to create robust, scalable, highly available and fault-tolerant solutions by learning different aspects of Solution architecture and next-generation architecture design in the Cloud environment.

Master erwin DM to deliver robust and precise designs for both operational and analytical projects. Steve and Jeff start from the basics, explaining data modeling concepts and how to get up and running with erwin DM (erwin DM). Through a hands-on approach, business analysts, data professionals, and project managers will learn step-by-step how to build effective conceptual, logical, and physical data models. Complete the stages in identifying essential business requirements, designing the logical data model, transposing those logical modeling objects into physical tables and columns, and even generating the implementation database scripts. This book contains seven parts. Part I provides a foundation in data modeling and Part II a foundation in erwin DM. Part III covers the design layer technique and its application using erwin DM, distinguishing conceptual, logical, physical, and operational data models. Part IV covers entities, domains, attributes, key groups, validation rules, default rules, and subject areas, along with how to implement them using erwin DM. Part V explains the physical data model and how to convert a logical data model to a physical data model in erwin DM. Become confident creating tables, columns, indexes, and views. Part VI reveals advanced features available within erwin DM, including user defined properties, naming standards, forward engineering, reverse engineering, complete compare, report designer, and the bulk editor. Part VII explains several important tools to use in combination with erwin DM, including erwin DM NoSQL, erwin Data Catalog, and erwin Data Literacy.

A quick and reliable way to build proven databases for core business functions Industry experts raved about The Data Model Resource Book when it was first published in March 1997 because it provided a simple, cost-effective way to design databases for core business functions. Len Silverston has now revised and updated the hugely successful First Edition, while adding a companion volume to take care of more specific requirements of different businesses. Each volume is accompanied by a CD-ROM, which is sold separately. Each CD-ROM provides powerful design templates discussed in the books in a ready-to-use electronic format, allowing companies and individuals to develop the databases they need at a fraction of the cost and a third of the time it would take to build them from scratch. With each business function boasting its own directory, this CD-ROM provides a variety of data models for specific implementations in such areas as financial services, insurance, retail, healthcare, universities, and telecom.

Essential Skills--Made Easy! Learn how to create data models that allow complex data to be analyzed, manipulated, extracted, and reported upon accurately. Data Modeling: A Beginner's Guide teaches you techniques for gathering business requirements and using them to produce conceptual, logical, and physical database designs. You'll get details on Unified Modeling Language (UML), normalization, incorporating business rules, handling temporal data, and analytical database design. The methods presented in this fast-paced tutorial are applicable to any database management system, regardless of vendor. Designed for Easy Learning Key Skills & Concepts--Chapter-opening lists of specific skills covered in the chapter Ask the expert--Q&A sections filled with bonus information and helpful tips Try This--Hands-on exercises that show you how to apply your skills Notes--Extra information related to the topic being covered Self Tests--Chapter-ending quizzes to test your knowledge Andy Oppel has taught database technology for the University of California Extension for more than 25 years. He is the author of Databases Demystified, SQL Demystified, and Databases: A Beginner's Guide, and the co-author of SQL: A Beginner's Guide, Third Edition, and SQL: The Complete Reference, Third Edition.

In this IBM Redbooks publication we describe and demonstrate dimensional data modeling techniques and technology, specifically focused on business intelligence and data warehousing. It is to help the reader understand how to design, maintain, and use a dimensional model for data warehousing that can provide the data access and performance required for business intelligence. Business intelligence is comprised of a data warehousing infrastructure, and a query, analysis, and reporting environment. Here we focus on the data warehousing infrastructure. But only a specific element of it, the data model - which we consider the base building block of the data warehouse. Or, more precisely, the topic of data modeling and its impact on the business and business applications. The objective is not to provide a treatise on dimensional modeling techniques, but to focus at a more practical level. There is technical content for designing and maintaining such an environment, but also business content. For example, we use case studies to demonstrate how dimensional modeling can impact the business intelligence requirements for your business initiatives. In addition, we provide a detailed discussion on the query aspects of BI and data modeling. For example, we discuss query optimization and how you can determine performance of the data model prior to implementation. You need a solid base for your data warehousing infrastructure . . . a solid data model.

[Copyright: e414049d64440059f44030f0803b8afe](https://www.ibm.com/press/antenna/2014/04/0414049d64440059f44030f0803b8afe)