



# A Manager's Guide to Data Warehousing



Laura L. Reeves

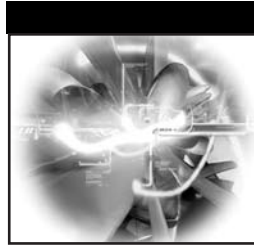




---

# **A Manager's Guide to Data Warehousing**





# **A Manager's Guide to Data Warehousing**

---

Laura L. Reeves



WILEY

Wiley Publishing, Inc.

Published by  
**Wiley Publishing, Inc.**  
10475 Crosspoint Boulevard  
Indianapolis, IN 46256  
www.wiley.com

Copyright © 2009 by Wiley Publishing, Inc., Indianapolis, Indiana

Published simultaneously in Canada

ISBN: 978-0-470-17638-2

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at [www.wiley.com/go/permissions](http://www.wiley.com/go/permissions).

**Limit of Liability/Disclaimer of Warranty:** The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a particular purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies contained herein may not be suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Neither the publisher nor the author shall be liable for damages arising herefrom. The fact that an organization or Web site is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Web site may provide or recommendations it may make. Further, readers should be aware that Internet Web sites listed in this work may have changed or disappeared between when this work was written and when it is read.

For general information on our other products and services please contact our Customer Care Department within the United States at (877) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Library of Congress Cataloging-in-Publication Data

Reeves, Laura L.

A manager's guide to data warehousing / Laura L. Reeves.

p. cm.

Includes index.

ISBN 978-0-470-17638-2 (paper/website)

1. Data warehousing—Management. I. Title.

QA76.9.D37R44 2009

005.74068—dc22

2009007401

**Trademarks:** Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. All other trademarks are the property of their respective owners. Wiley Publishing, Inc. is not associated with any product or vendor mentioned in this book.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

---



## About the Author

Laura L. Reeves started designing and implementing data warehouse solutions in 1986. Since then she has been involved in hundreds of projects. She has extensive experience in end-to-end data warehouse development, including developing comprehensive project plans, collecting business requirements, developing business dimensional models, designing database schemas (both star and snowflake designs), and developing enterprise data warehouse architecture and strategies. These have been implemented for many business functions for private and public industry.

Laura co-founded StarSoft Solutions, Inc., in 1995 and has been a faculty member with The Data Warehousing Institute since 1997. She is a contributing author of *Building a Data Warehouse for Decision Support* (Prentice Hall, 1996) and a co-author of the first edition of *The Data Warehouse Lifecycle Toolkit* (Wiley, 1998). Laura graduated magna cum laude from Alma College with a bachelor of science degree in mathematics and computer science, with departmental honors.





---



# Credits

**Executive Editor**

Robert Elliott

**Development Editor**

Sara Shlaer

**Technical Editor**

Jonathan Geiger

**Production Editor**

Melissa Lopez

**Copy Editor**

Luann Rouff

**Editorial Manager**

Mary Beth Wakefield

**Production Manager**

Tim Tate

**Vice President and Executive  
Group Publisher**

Richard Swadley

**Vice President and Executive  
Publisher**

Barry Pruett

**Associate Publisher**

Jim Minatel

**Proofreader**

Josh Chase, Jen Larsen, and  
Kyle Schlesinger, WordOne

**Indexer**

Robert Swanson

**Cover Image**

© Digital Vision



---



# Acknowledgments

I have been very blessed with great family, friends, and colleagues. I would like to thank the many clients and colleagues who have challenged me, pushed me, and collaborated with me on so many initiatives over the years. I appreciate the opportunity to work with such high-quality people. I want to acknowledge the contributions that have been made to the data warehousing industry and to me personally by the amazing people who worked at Metaphor. I want to express my gratitude to my dear friend and colleague Paul Kautza for his belief in me and for all his hard work all these years.

Thanks are also due to the dedicated staff at Wiley who believed in me and had great patience to help see this project through. Thanks to Bob Elliott for being the impetus to get this project started and to Sara Shlaer and Rosanne Koneval for their detailed efforts to produce a quality product. I want to express appreciation to Cindi Howson for her insight on business intelligence tools.

I want to extend a sincere and special thank you to Jonathon Geiger for his meticulous comments and suggestions.

I also want to thank two very special people who have provided unflagging support and encouragement every step of the way: my friends Ingrid Korb and Paula Johnson. I am not sure I could have done this without you!

Of course, none of this would be possible without the dedication, sacrifice, love, and support given to me by my family: Mark, Ryan, Michael, and Leah.



---



# Contents

<b>Introduction</b>	<b>xxiii</b>
<b>Part One The Essentials of Data Warehousing</b>	<b>1</b>
<b>Chapter 1 Gaining Data Warehouse Success</b>	<b>3</b>
The Essentials of Data Warehousing	3
What Is a Data Warehouse?	4
Differences Between Operational and DW Systems	4
The Data Warehousing Environment	4
What Is a Data Model?	6
Understanding Industry Perspectives	7
Design and Development Sequence	8
Why Build a Data Warehouse?	11
The Value of Data Warehousing	12
The Promises of Data Warehousing	15
Keys to Success	16
Developing and Maintaining Strong Business and Technology Partnerships	17
Identifying True Business Requirements	17
Shifting to a Global Perspective	18
Overcoming Unrealistic Expectations	19
Providing Clear Communication	20
Treating Data As a Corporate Asset	21
Effectively Leveraging Technology	21
Roadblocks to Success	22
Believing the Myth: “If You Build It, They Will Come”	22
Falling into the Project Deadline Trap	23

Failing to Uphold Organizational Discipline	23
Lacking Business Process Change	24
Narrowing the Focus Too Much	25
Resting on Your Laurels	27
Relying on the Technology Fix	27
Getting the Right People Involved	28
Finding Lost Institutional Knowledge	29
Summary	30
<b>Chapter 2 The Executive's FAQ for Data Warehousing</b>	<b>31</b>
Question: What is the business benefit of a data warehouse?	32
Answer	32
Question: How much will it cost?	33
Answer	33
Question: How long will it take?	34
Answer	35
Question: How can I ensure success?	36
Answer	36
Question: Do other companies really build these in 90 days?	37
Answer	37
Question: How will we know we are doing this right?	38
Answer	38
Question: Why didn't this work last time? What is different this time?	39
Answer	39
Question: Do we have the right technology in place?	39
Answer	40
Question: Are we the only company with data warehouse problems?	40
Answer	41
Question: Will I get one version of the truth?	41
Answer	42
Question: Why can't we just use our current systems?	43
Answer	44
Question: Will the data warehouse replace our old systems?	45
Question: Who needs to be involved?	45
Question: Do we know where we are going? How will we know when we get there?	46
Answer	46
Question: How do we get started and stay focused?	47
Answer	47
Summary	48

<b>Part Two</b>	<b>The Business Side of Data Warehousing</b>	<b>49</b>
<b>Chapter 3</b>	<b>Understanding Where You Are and Finding Your Way</b>	<b>51</b>
	Assessing Your Current State	51
	What Is Your Company's Strategic Direction?	52
	What Are the Company's Top Initiatives?	54
	How Healthy Is Your Data?	55
	Does the Business Place Value on Analysis?	56
	Reflecting on Your Data Warehouse History	57
	Understanding Your Existing Reporting Environment	58
	Finding the Reporting Systems	59
	Compiling an Inventory	60
	Identifying the Business Purpose	61
	Discovering the Data You Already Have	63
	Understanding the People	65
	Tracking Technology and Tools	65
	Understanding Enterprise Resources	66
	Netting It All Out	68
	Introducing the Case Studies	70
	The Call Center Data Warehouse Project	70
	In Real Life	70
	Giant Company	71
	Agile, Inc.	72
	Summary	72
<b>Chapter 4</b>	<b>Successful IT–Business Partnerships</b>	<b>75</b>
	What a Partnership Really Means	75
	What the Business Partners Should Expect to Do	76
	Business Executives and Senior Management	78
	The Executive Business Sponsor	78
	Business Managers	81
	The Business Champion	82
	Business Analysts	83
	Helping the Business Analyst Deal with Change	85
	Business User Audience	86
	Project Manager	86
	What You Should Expect from IT	88
	CIO/IT Executive Sponsor	89
	Data Warehouse Manager	89
	Business Systems Analyst	90
	Source System Analyst	91
	Data Modeler/Data Architect	92

ETL Developer(s)	93
Business Intelligence Application Developer	94
Other Supporting Roles	95
Tips for Building and Sustaining a Partnership	95
Leveraging External Consulting	97
Building Strong Project Teams	98
Effective Communication	99
Netting Out Key Messages	99
Presenting in Business Terms	100
Meeting Preparation	101
Presentation Tips	102
When to Communicate	103
Partnerships Beyond a Project	104
The Decision-Making Process	104
Executive Steering Committee	104
DW Business Support Team	106
Enterprise Considerations	107
In Real Life	107
A Glimpse into Giant, Co.	107
Insight from Agile, Inc.	108
Summary	109
<b>Chapter 5 Setting Up a Successful Project</b>	<b>111</b>
Defining the Project	111
Setting Up the Project Charter	112
Documenting Project Scope	117
Developing a Statement of Work	117
How Much Will It Cost?	120
Project Approval	122
Starting the Project	122
Launching the Project	123
Managing a Successful Project	124
Issue Tracking	124
Using Project Change Control	125
Discussing Change in Business Terms	126
Managing Expectations	128
In Real Life	129
Structured Projects with Giant	129
Freedom for Creativity at Agile, Inc.	130
Summary	131
<b>Chapter 6 Providing Business Requirements</b>	<b>133</b>
What Requirements Are Needed?	134
Peeling Back the Layers of Requirements Gathering	134



Who Provides Input?	137
Who Gathers the Requirements?	137
Providing Business Requirements	138
Strategic Requirements	138
Broad Business Requirements	140
Business Analyses	143
Business Data Requirements	145
Systems and Technical Requirements	147
Communicating What You Really Need	149
What Else Would Help the Project Team?	150
Data Integration Challenges	151
Assess Organizational Motivation	151
Complete Picture of the Data	152
What If No One Is Asking?	152
Practical Techniques for Gathering Requirements	153
Interview Session Characteristics	153
Individual Interviews	153
Group Interviews	153
Project Team Participation	154
Interview Tips	154
Who Needs to Be Included?	155
Setting a Good Example	156
Preparing for Interview Sessions	157
Conducting the Interview Sessions	157
Capturing Content: Notes vs. Tapes	157
Running the Interview	158
Concluding the Interview	158
Putting the Pieces Together	158
Individual Interview Documentation	159
Responsibilities	159
Business Themes	159
Business Data	160
Consolidated Requirements Documentation	161
Executive Summary	161
Consolidated Business Themes	162
Candidate Business Analyses	162
Consolidated Business Data Requirements	162
Identification of Non-Data Warehouse Requirements	163
Common Requirements Gathering Challenges	163
Sifting Through Reports	163
Listing Data Elements	164
Developing Functional Specifications	164
Moving Beyond Immediate	164
Lack of Requirements	165
The Cynic	165

Setting Attainable Goals	166
Exploring Alternatives	167
Setting Priorities	168
In Real Life	170
A Glimpse into Giant Company	170
Insight from Agile, Inc.	170
Summary	171
<b>Part Three Dealing with the Data</b>	<b>173</b>
<b>Chapter 7 Modeling the Data for your Business</b>	<b>175</b>
The Purpose of Dimensional Models	176
Ease of Use	176
Query Performance	177
Understanding Your Data	177
What Is a Dimensional Model?	178
Dimensions	178
Facts	180
Using Both Parts of the Model	180
Implementing a Dimensional Model	181
Diagramming Your Dimensional Model	182
The Business Dimensional Model	182
Business Dimensions	183
Fact Groups	184
A Call Center Case Study	186
Call Center Dimensions	187
Date Dimension	187
Time Dimension	187
Customer Dimension	189
Employee Dimension	191
Call Dimension	191
Call Outcome Dimension	194
Employee Task Dimension	195
Call Center Fact Groups	196
Calls Fact Group	196
Call Center Time Tracking Fact Group	196
Call Forecast Fact Group	198
Working with the Model	199
Business Dimensional Model Index	200
Enterprise Considerations	200
Conformed Dimensions	200
Conformed Facts	202
Practical Guidelines	202
Guidelines for a Single Dimension	202

Guidelines for a Single Fact Group	203
Characteristics of the Model across the Enterprise	204
Business Participation in the Modeling Process	205
Creating the First Draft	205
Preparing for Modeling Sessions	205
Brainstorming the Framework	206
Drafting the Initial Dimensions	206
Drafting the Initial Fact Groups	207
Documenting the Model	208
Logging Questions and Issues	208
Building the Business Measures Worksheet	209
Preliminary Source to Target Data Map	211
Completing or Fleshing Out the Model	211
Working Through the Issues	211
Completing the Documentation	212
Working Through All the Data Elements	212
Refining the Model	213
Business Reviews of the Model	213
Small Business Reviews	214
When Are You Done?	214
Gaining Final Commitment	215
Expanding Business Data Over Time	215
Enhancing Dimensions	215
Adding More Fact Groups	215
Reflecting on Business Realities: Advanced Concepts	216
Supporting Multiple Perspectives: Multiple Hierarchies	216
Tracking Changes in the Dimension: Slowly Changing Dimensions	216
Depicting the Existence of a Relationship: Factless Fact Tables	218
Linking Parts of a Transaction: Degenerate Dimensions	219
Pulling Together Components: Junk Dimensions	221
Multiple Instances of a Dimension: Role Playing	222
Other Notation	224
Dimension Connectors	224
Clusters of Future Attributes	225
Notation Summary	225
Taking the Model Forward	225
Translating the Business Dimensional Model	226
Dimension Table Design	226
Translating Fact Groups	227
Physical Database Design	228
In Real Life	228
A Glimpse into Giant Co.	229
Insight from Agile, Inc.	229
Summary	230

<b>Chapter 8</b>	<b>Managing Data As a Corporate Asset</b>	<b>231</b>
	What Is Information Management?	232
	Information Management Example—Customer Data	235
	IM Beyond the Data Warehouse	239
	Master Data Management	240
	Master Data Feeds the Data Warehouse	242
	Finding the Right Resources	242
	Data Governance	243
	Data Ownership	243
	Who Really Owns the Data?	244
	Your Responsibilities If You Are “the Owner”	246
	What are IT’s Responsibilities?	247
	Challenges with Data Ownership	247
	Data Quality	248
	Profiling the Data	249
	How Clean Does the Data Really Need to Be?	250
	Measuring Quality	250
	Quality of Historical Data	251
	Cleansing at the Source	253
	Cleaning Up for Reporting	254
	Managing the Integrity of Data Integration	254
	Quality Improves When It Matters	256
	Example: Data Quality and Grocery Checkout Scanners	257
	Example: Data Quality and the Evaluation of Public Education	257
	Realizing the Value of Data Quality	258
	Implementing a Data Dictionary	259
	The Data Dictionary Application	259
	Populating the Data Dictionary	261
	Accessing the Data Dictionary	263
	Maintaining the Data Dictionary	263
	Getting Started with Information Management	264
	Understanding Your Current Data Environment	264
	What Data Do You Have?	265
	What Already Exists?	266
	Where Do You Want to Be?	267
	Develop a Realistic Strategy	268
	Sharing the Information Management Strategy	269
	Setting Up a Sustainable Process	270
	Enterprise Commitment	270
	The Data Governance Committee	270
	Revising the Strategy	271

In Real Life	271
A Glimpse into Giant, Co.	272
Insight from Agile, Inc.	272
Summary	274
<b>Part Four    Building the Project</b>	<b>275</b>
<b>Chapter 9    Architecture, Infrastructure, and Tools</b>	<b>277</b>
What Is Architecture?	278
Why Do We Need Architecture?	278
Making Architecture Work	281
Data Architecture	282
Revisiting DW Goals	283
Components of DW Data Architecture	285
A Closer Look at Common Data Warehouse Architectures	286
Bottom-Up Data Architecture	286
Top-Down Data Architecture	290
Publish the Data: Data Marts	294
Adopting an Architecture	295
Technical Architecture	297
Technical Architecture Basics	298
Components of Technical Architecture	299
Infrastructure	300
Technical Architecture in Action	300
What You Need to Know about Technical Architecture	301
Navigating the Technology Jungle	302
Weighing Technology Options	303
Best of Breed	303
End-to-End Solutions	303
Deciding Not to Buy a Tool	304
Finding the Right Products	304
Requests for Information or Proposals	305
Business Participation in the Selection Process	305
Understanding Product Genealogy	306
Understanding Value and Evaluating Your Options	306
Cutting through the Marketing Hype	308
The Value of References	309
Making Architecture Work for You	310
Just-In-Time Architecture	311
In Real Life	311
Architecture at Giant	311
Agile Ignores the Need for Architecture	312
Summary	313

<b>Chapter 10 Implementation: Building the Database</b>	<b>315</b>
Extract, Transform, and Load (ETL) Fundamentals	315
What Work Is Being Done?	315
ETL System Functionality	317
Extraction	318
Transformation	318
Load	322
The Business Role in ETL	323
Why Does the Business Need to Help?	323
Defining Business Rules	324
Defining Expected Results—The Test Plan	325
Development Support	326
Testing the ETL System—Is the data Right?	326
Why Does It Take So Long and Cost So Much?	327
Balancing Requirements and Data Reality	329
Discovering the Flaws in Your Current Systems	330
Applying New Business Rules	331
Working Toward Long-Term Solutions	332
Manually Including Business Data	333
Tracking Progress—Are We There Yet?	333
What Else Can You Do to Help?	334
Encouragement and Support	334
Ensuring Continued Business Participation	335
Proactive Communication	336
In Real Life	337
Building the Data Warehouse at Giant, Co.	337
Agile, Inc., Builds a Data Warehouse Quickly	338
Summary	339
<b>Chapter 11 Data Delivery: What you Finally See</b>	<b>341</b>
What Is Business Intelligence?	341
Business Intelligence without a DW	342
BI in Action	343
Tabular Reports	343
Parameter-Driven Reports	343
Interactive Reports—Drilling Down and Across	344
Exception Reports	344
Other BI Capabilities	345
Complex Analysis	345
BI Building Blocks	346
Data Content—Understanding What You Have	346
Navigation—Finding What You Need	347

Presentation—How Do You Want to See Results?	347
Delivery—How Do You Receive the Results?	351
Supporting Different Levels of Use	352
Construction of the BI Solution	354
Planning for Business Change	354
Design—What Needs to Be Delivered?	355
Development	357
Testing BI Applications and Validating Data	358
Additional Responsibilities	359
Security—Who Can Look at the Data?	359
System Controls—Who Can Change What?	360
Planning a Successful Launch	361
Marketing the Solution	361
Learning to Use the Data without a Technical Degree	362
Learning about the Data	362
Learning about the BI Tool/Application	362
Ensuring That the Right Help Is Available	363
In Real Life	364
BI at Giant Company	364
Agile, Inc. Dives into BI	365
Summary	366
<b>Part Five</b>	
<b>Next Steps—Expanding On Success</b>	<b>367</b>
<b>Chapter 12</b>	
<b>Managing the Production Data Warehouse</b>	<b>369</b>
Finishing the Project	369
Recapping the BI Application Launch	369
Post-Implementation Review	370
Looking Back—Did you Accomplish Your Objectives?	371
Adopting the Solution	371
Tracking Data Warehouse Use	372
Getting the Rest of the Business Community on Board	372
Business Process Change	374
Changing How Data Is Used	374
Streamlining Business Processes	374
Encouraging Change	375
The Production Data Warehouse	375
Staffing Production Activities	376
Maintaining the Environment	376
Keeping Up with Technology	376
Monitoring Performance and Capacity Planning	378
Maintaining the Data Warehouse	380
Maintaining the ETL System	380
Maintaining the BI Application	381

Tracking Questions and Problems	382
Fixing Bugs	384
When the Data Warehouse Falls Short	384
Common Causes for a Stalled Warehouse	385
Jump-Starting a Stalled Data Warehouse	388
Conducting an Assessment	388
Determining What Can Be Salvaged	389
Developing a Plan to Move On	390
Aligning DW Objectives with Business Goals	391
Getting It Right This Time	392
Launching the Improved Data Warehouse and BI Solution	393
In Real Life	394
Lack of Support for the Production DW at Giant Co.	394
Unleashing BI at Agile, Inc.	395
Summary	396
<b>Chapter 13 Achieving Long-Term Success</b>	<b>397</b>
Planning for Expansion and Growth	397
Exploring Expansion Opportunities	398
Prioritization of Feedback	399
Managing Enterprise DW Resources	400
Creating an Enterprise Data Warehouse Team	400
The Centralized Enterprise Data Warehouse Team	401
The Virtual Enterprise Data Warehouse Team	401
Enterprise DW Team Responsibilities	403
Funding the Enterprise DW Team	404
Pushing into the Future	405
Embedded Business Intelligence	405
Operational Business Intelligence	406
Real-Time Data Warehousing	407
Unstructured Data	408
Monitoring Industry Innovation	409
Moving Toward Business Value	410
Measuring Success One Step at a Time	410
Adjusting Expectations to Reality	412
Keeping the Momentum Going	413
Celebrating Progress	416
Success Can Be Attained	417
Conclusion	419
<b>Glossary</b>	<b>421</b>
<b>Index</b>	<b>429</b>



- [Undisputed: How to Become World Champion in 1,372 Easy Steps book](#)
- [Renegade of Kregen \(Dray Prescot, Book 13; Krozair Cycle, Book 2\) online](#)
- [download The Origins of American Social Science \(Ideas in Context\)](#)
- [Uncle John's Bathroom Reader History's Lists pdf](#)
- [France, a Nation on the Verge of a Nervous Breakdown pdf, azw \(kindle\), epub](#)
  
- <http://betsy.wesleychapelcomputerrepair.com/library/Undisputed--How-to-Become-World-Champion-in-1-372-Easy-Steps.pdf>
- <http://junkrobots.com/ebooks/Sylvia--A-Novel.pdf>
- <http://cavalldecartro.highlandagency.es/library/The-Origins-of-American-Social-Science--Ideas-in-Context-.pdf>
- <http://junkrobots.com/ebooks/Uncle-John-s-Bathroom-Reader-History-s-Lists.pdf>
- <http://www.experienceolvera.co.uk/library/France--a-Nation-on-the-Verge-of-a-Nervous-Breakdown.pdf>