The Data Warehouse Lifecycle Toolkit

Second Edition

Practical Techniques for Building Data Warehouse and Business Intelligence Systems

Ralph Kimball
Margy Ross
Warren Thornthwaite
Joy Mundy
Bob Becker
The Data Warehouse Lifecycle Toolkit

Second Edition
About the Authors

The authors’ professional careers have followed remarkably similar paths. Each author has focused on data warehousing and business intelligence (DW/BI) consulting and education for more than fifteen years. Most worked together at Metaphor Computer Systems, a pioneering decision support vendor, in the 1980s. All the authors are members of the Kimball Group and teach for Kimball University. They contribute regularly to Intelligent Enterprise magazine and other industry publications; most have previously written books in the Toolkit series.

Ralph Kimball founded the Kimball Group. Since the mid 1980s, he has been the DW/BI industry’s thought leader on the dimensional approach and trained more than 10,000 IT professionals. Ralph has his Ph.D. in Electrical Engineering from Stanford University.

Margy Ross is President of the Kimball Group. She has focused exclusively on DW/BI since 1982 with an emphasis on business requirements analysis and dimensional modeling. Margy graduated with a BS in Industrial Engineering from Northwestern University.

Warren Thornthwaite began his DW/BI career in 1980. After managing Metaphor’s consulting organization, he worked for Stanford University and WebTV. Warren holds a BA in Communications Studies from the University of Michigan and an MBA from the University of Pennsylvania’s Wharton School.

Joy Mundy has focused on DW/BI systems since 1992 with stints at Stanford, Web TV, and Microsoft’s SQL Server product development organization. Joy graduated from Tufts University with a BA in Economics, and from Stanford University with an MS in Engineering Economic Systems.

Bob Becker has helped clients across a variety of industries with their DW/BI challenges and solutions since 1989, including extensive work with health care organizations. Bob has a BSB in Marketing from the University of Minnesota’s School of Business.
Acknowledgments

First, thanks to our students, clients, readers, and former colleagues for supporting, teaching, and influencing us. One of the authors recently received a fortune cookie that read, “You learn most when teaching others.” We couldn’t agree more. Our Kimball University students have pushed us to provide precise, specific guidance and kept us on our toes with their questions. Similarly, the challenges faced by our Kimball Group consulting clients have become our challenges, and have kept us grounded in reality. Finally, ex-colleagues have contributed to our thinking about the concepts in this book, including Laura Reeves who participated as a co-author of the first edition of the Lifecycle Toolkit.

Beginning with our associates from the early days at Metaphor, through Red Brick, Stanford University, DecisionWorks Consulting, InfoDynamics, and Microsoft, we’ve learned lots from each of you.

Thanks to the Wiley team for making this book a reality. Bob Elliott’s subtle, yet persistent prodding got the project off the ground. Sara Shlaer did a wonderful job editing our text with an incredible amount of patience, tenacity, and attention to detail. Deb Banninger and the behind-the-scenes folks worked tirelessly to deliver a quality product. We’ve enjoyed working with all of you.

Finally, thanks to our spouses, partners, and children for putting up with the demands of our careers, while supporting us unconditionally. You’ve suffered through late nights and missed vacations alongside us. Thanks to Julie Kimball, Sara Kimball Smith, and Brian Kimball, Scott and Katie Ross, Elizabeth Wright, Tony Navarrete, and Pam, Elisa, and Jenna Becker. We couldn’t have done it without you!
## Contents at a Glance

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td>Introducing the Kimball Lifecycle</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>Launching and Managing the Project/Program</td>
<td>15</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Collecting the Requirements</td>
<td>63</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>Introducing the Technical Architecture</td>
<td>109</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>Creating the Architecture Plan and Selecting Products</td>
<td>179</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>Introducing Dimensional Modeling</td>
<td>233</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>Designing the Dimensional Model</td>
<td>287</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>Designing the Physical Database and Planning for Performance</td>
<td>327</td>
</tr>
<tr>
<td>Chapter 9</td>
<td>Introducing Extract, Transformation, and Load</td>
<td>369</td>
</tr>
<tr>
<td>Chapter 10</td>
<td>Designing and Developing the ETL System</td>
<td>425</td>
</tr>
<tr>
<td>Chapter 11</td>
<td>Introducing Business Intelligence Applications</td>
<td>473</td>
</tr>
<tr>
<td>Chapter 12</td>
<td>Designing and Developing Business Intelligence Applications</td>
<td>505</td>
</tr>
<tr>
<td>Chapter 13</td>
<td>Deploying and Supporting the DW/BI System</td>
<td>541</td>
</tr>
<tr>
<td>Chapter 14</td>
<td>Expanding the DW/BI System</td>
<td>579</td>
</tr>
</tbody>
</table>
Contents

Acknowledgments ix

Introduction xxxi

Chapter 1 Introducing the Kimball Lifecycle 1
Lifecycle History Lesson 1
Lifecycle Milestones 3
Program/Project Planning 4
Program/Project Management 4
Business Requirements Definition 5
Technology Track 5
Technical Architecture Design 5
Product Selection and Installation 6
Data Track 6
Dimensional Modeling 6
Physical Design 6
ETL Design and Development 7
Business Intelligence Application Track 7
BI Application Design 7
BI Application Development 7
Deployment 7
Maintenance 8
Growth 8
Using the Lifecycle Roadmap 8
Lifecycle Navigation Aids 9
Lifecycle Vocabulary Primer 9
Data Warehouse versus Business Intelligence 10
ETL System 11
Business Process Dimensional Model 12
Business Intelligence Applications 13
Conclusion 14