



Mastering Data Modeling

*A User-Driven
Approach*



John Carlis
Joseph Maguire

Praise for *Mastering Data Modeling: A User-Driven Approach*

This is without doubt the most complete and worthy “common sense” approach to data modeling that I have read. It spans the ‘stupidity wars’ of differences of notation and nomenclature with a clear and easy elegance that espouses common sense over dogma. . . . An excellent work which deserves to be a foundation of every programmer and software designer’s bookshelf. Without understanding of data, there can be no intelligent software design.

Paul Irvine
Senior Manager
Emerald Solutions

I will also be able to use this book to make myself not only a better data modeler and database designer, but a better overall systems analyst, project planner, and business communicator. A second (and perhaps third) reading is definitely in order!

Matthew Keranen
Consulting Database Administrator

An invaluable resource for anyone who wants to learn about data modeling, this book is exceptionally powerful because its developed from real world experiences—it is a pragmatic guide for data modeling practitioners.

Peter O’Kelley
Senior Consultant/Analyst
The Patricia Seybolt Group

This book presents a detailed, structured approach to the process of data modeling with emphasis on interacting with users, a unique feature which impressed me.

David McGoveran
Alternative Technologies, Inc.

I hope the book is well received. It deserves to be, as it provides extremely clear, practical thinking on an important but often misunderstood subject.

Mitch Kapur
Founder of the Lotus Development Corporation

Mastering Data Modeling: A User-Driven Approach

Table of Contents

Contents

Foreword

Preface

Chapter 1 Introduction

- Logical Data Structures and Physical Data Storage

 - Logical: Thinking About WHAT Without HOW

 - Data: Thinking About Data Without Processing

 - Structure: Articulating Types Without Instances

 - Example

 - Extending the Example LDS: Shifting the Burden

- Summary

- Exercises

Chapter 2 Good Habits

- Employ the Users Language and Vocabulary

- Be Rigorous

- Dont Rely on the Opinion of a Single Expert; Ask Several

- Ask First About Data, Not About Processing

- Master the Shapes of Data

- Use a Notation That Helps You Realize These Good Habits

- Summary

- Exercises

Table of Contents

Chapter 3 Reading an LDS with Sentences

Sentences About What Users Can Remember

Sentences About Words Inside Boxes

Sentences About (Unlabeled) Box-to-Box Lines

Chicken Feet

Similarities Between Sentences

Sentences About Labeled Box-to-Box Lines

Labels and Similarities Between Sentences

Similarities Between Sentences Revisited

Sentences About Differentiating Things from Each Other

One-Bar Boxes

Multiple-Bar Boxes

Outside-the-Box Bars

A Shorthand

Another Common Error You Should Avoid

Sentences You Should Not Say

Some Complete Examples

Summary

Exercises

Chapter 4 Vocabulary of LDS

Vocabulary Overview

Entity and Entity Instance

Attribute and Attribute Value

Relationship and Relationship Instance

Link and Link Value

Maximum Degree

Descriptor and Descriptor Value

Identifier and Identifier Value

Table of Contents

Identifying Descriptor

Degree-One Descriptor

Degree-Many Descriptor

One-Many Relationship

Many-Many Relationship

One-One Relationship

To-be Relationship

Not-to-be Relationship

Described Entity and Describing Entity

Tiebreaker

A Bit More About Entities, Attributes, and Relationships

Entities versus Attributes

Entities versus Relationships

LDS Reading Rules Revisited

Responsibility for Speaking Well

Summary (and a Chance to Check Your Progress)

Using a Plural Noun for the Described Entity

Overlooking Links

Overlooking the Conventional Role of Identifying Descriptors

Confusing Identifiers with Identifying Descriptors

Overlooking the Similarity Between an Identifying Attribute and an Identifying Link

Neglecting to Read Both Links of a Relationship

Exercises

Chapter 5 Visualizing Allowed and Disallowed Instances

Show the Data and Say Something About It

Plan Your Notes by Considering Elemental Parts of the LDS

Some Notes Apply Simply Because the Diagram Includes an Attribute

Some Notes Apply Because the Diagram Includes a Degree-One Link

Table of Contents

Some Notes Apply Because the Diagram Includes a Degree-Many Link

Some Notes Apply Because the Diagram Includes a Single-Descriptor Identifier

Some Notes Apply Because the Diagram Includes a Multiple-Descriptor Identifier

If an Identifier Has Three or More Descriptors, the Notes Can Be More Elaborate Than the Notes for a Two-Descriptor Identifier

As You Visualize Data, Dont Lose Sight of the Goal

More Data Is Not Necessarily Better

Less Data Is Not Necessarily Better

Dont Limit Yourself to a Single Way of Visualizing the Data

Exercises

Chapter 6 A Conversation with Users About Creatures and Skills

Summary

Exercises

Story Interlude

Chapter 7 Introduction to Mastering Shapes

Definition of Shape

Mastering Shapes

Reading a Shape Aloud in Several Ways

Visualizing Sample Data in Several Formats

Discussing and Illustrating Noteworthy Disallowed Data

Finding and Focusing on Shapes Within a Large LDS

Recognizing the Differences Between Shapes That Are Similar but Not Identical

Recognizing the Similarity Between Seemingly Dissimilar Shapes

Distinguishing Between Legitimate Shapes and Syntactically

Table of Contents

Invalid LDS Fragments

Knowing How Shapes Are Likely to Evolve

Asking Questions That Help Users Choose Between Two Similar
Shapes

Knowing When to Ask Questions of Users

Knowing When and How to Modify the LDS to Make a Shape Evolve

Understanding the Relative Frequency of the Various Shapes

Referring to each Fundamental Shape by Its Name

Exercises

Chapter 8 One-Entity, No-Relationship Shapes

Shape: Common Independent Entity

Shape: Lonely-Attribute Independent Entity

Shape: Aggregate Independent Entity

Shape: Dependent Entity

Shape: Intersection Entity

Shape: Subordinate Entity

Shape: One-Many Collection Entity

Shape: Many-Many Collection Entity

Unnamed Possibilities

Exercises

Chapter 9 One-Attribute Shapes

Scale

Shape: Nominal-Scale Attributes

Shape: Numeric-Scale Attributes

Shape: Ordinal-Scale Attributes

Shape: Boolean-Scale Attributes

Table of Contents

Scale and Datatype

Scale and Attribute Names

Fine Distinctions of Scale

Scale and Abstract Datatypes

Summary of How Scale Restricts an Attribute

Exercises

Chapter 10 Two-Entity Shapes

Two Entities, One Relationship

One-Many Shapes

Shape: Plain One-Many Relationship

Shape: One-Many Relationship Making a Dependent or Intersection Entity

Shape: One-Many Relationship Making a Collection Entity

One-One Shapes

Shape: Plain One-One Relationship

Shape: One-One Relationship Making a Subordinate Entity

Shape: To-be One-One Relationship

Shape: Not-to-be One-One Relationship

Shape: To-be One-One Relationship Making a Subordinate Entity

Shape: Plain To-be One-One Relationship

Shape: Plain Not-to-be One-One Relationship

Shape: Not-to-be One-One Relationship Making a Subordinate Entity

Thinking About One-One Relationship Shapes

Many-Many Shapes

Shape: Plain Many-Many Relationship

Shape: Many-Many Relationship Making a Collection Entity

Two Entities, Two Relationships

Shape: Two One-One Relationships

One-One and One-Many Relationship

Table of Contents

Shape: Not-to-be Relationship and One-Many Relationship

Shape: To-be Relationship and One-Many Relationship

Two One-Many Relationships

Shape: Two Same-Direction One-Many Relationships

Shape: Two Opposite-Direction One-Many Relationships

Shape: Many-Many Relationship Plus Some Other Relationships

Shape: Two Entities, n Relationships

Exercises

Chapter 11 Shapes with More Than Two Entities

Shape: Chicken Feet In

Shape: Chicken Feet Out

Shape: Chicken Feet Across

Shape: Subordinates Out

Shape: Subordinates Across

Shape: Multiple Plain To-be Relationships

Shape: Multiple To-be Relationships

Shape: Multiple Short Paths

Exercises

Chapter 12 Shapes with Reflexive Relationships

Shape: One-One Reflexive Relationship

Sequence Data and Cyclic Sequence Data

Ordered Pairs

Shape: One-Many Reflexive Relationship

Shape: Many-Many Reflexive Relationship

Exercises

Story Interlude

Chapter 13 LDS Syntax Rules

Table of Contents

Within Any LDS, Each Entity, Attribute, Relationship, and Link Has an
Official Name That Is Unique

No Reflexive Relationship Is a To-be Relationship

Between Any Pair of Entities, There Is at Most One To-be Relationship

Each Entity Has at Least One Identifier

An Entity Can Have Several Identifiers

No Identifier Can Be a Strict Subset of Another

The LDS Cannot Contain Any Cycles of Identification Dependency

No Link of a Reflexive Relationship Can Contribute to an Identifier

Both Links of a Relationship Cannot Contribute to Identifiers

A Single-Descriptor Identifier Cannot Include the Degree-One Link of a
One-Many Relationship

A Multiple-Descriptor Identifier Cannot Include a Link of a One-One
Relationship

A Multiple-Descriptor Identifier Cannot Include the Degree-Many Link of
a One-Many Relationship

A Relationship Has Either Two Labels or Zero Labels

All One-One Relationships Have Labels

All Reflexive Relationships Have Labels

Between Any Pair of Entities, There Is at Most One Unlabeled
Relationship

Valid Relationships

Exercises

Chapter 14 Getting the Names Right

Entity Names

Too-Exclusive and Too-Inclusive Entity Names

Too-Coarse Entity Names

Table of Contents

Too-Fine Entity Names

Completely Inaccurate Names

Overloaded Names

Working with Users to Get the Entity Names Right

Expect to Work Hard on Naming Entities

Be Willing to Work Hard Because Its Worth It

Manage the Difficulty: Choose the Right Moments to Work Hard

Manage the Diffi]culty: Use the Expertise at Your Disposal

Manage the Difficulty: Teach the Users a Helpful Basic Principle

Example: Checking a Name That You Suspect is Too Coarse

Naming Attributes

Naming Relationships and Links

Exercises

Chapter 15 Official Names

Official Names Can Be Awkward

Coping with Awkwardness in Official Names of Links

Coping with Awkwardness in Official Names of Relationships

A Few Notes About Official Names and To-be Relationships

Exercise

Chapter 16 Labeling Links

Exercises

Chapter 17 Documenting an LDS

The Audience

Front Matter

Entity Documentation

Attribute Documentation

Link Documentation

Table of Contents

Relationship Documentation

Fragment Documentation

Constraint Documentation

Issues List

Supplemental Material for Secondary Audiences

 Software Professionals

 Business Designers

Exercise

Story Interlude

Chapter 18 Script for Controlled Evolution: The Flow

 Script for The Flow

 Discussing a Not-to-be Relationship

 Flow Investigation: Discover Relationship

 Listening to Users

 Discovering Hidden Relationships

 Explicitly Asking for Relationships

 Securing an Answer to the Question

 Possible Answers to the Question

 Flow Stage: Not-to-be Relationship

 Flow Investigation: Seek a Chicken Foot

 Securing an Answer to the Question

 Possible Answers to the Question

 Flow Investigation: Seek a One-Many Relationship

 Securing an Answer to the Question

 Possible Answers to the Question

 Flow Investigation: Seek a Many-Many Relationship

 Securing an Answer to the Question

 Possible Answers to the Question

Table of Contents

Flow Stage: One-One, Not-to-be Relationship

Flow Stage: One-Many Relationship

Flow Stages: Initial Many-Many Relationship and New Intersection Entity

Developing a Chicken-Feet-In Shape

Preview of Working on a New Intersection Entity

Flow Investigation: Seek Descriptors for Intersection Entity

Flow Investigation: Seek Tiebreaker

Flow Investigation: Consider Overidentification

Flow Investigation: Seek Independent Entity

Discussing a To-be Relationship

Flow Investigation: Consider Synonymy

Securing an Answer to the Question

Possible Answers to the Question

Flow Investigation: Consider Subordination

Continuing the Discussion

Flow Continuation: Seek Other Relationships

Flow Continuation: Seek Further Evolution for a One-Many
Relationship

Flow Continuation: Seek Further Evolution for the
Chicken-Feet-Across Shape

Flow Continuation: Seek Further Evolution for the Chicken-Feet-In Shape

Exercises

Chapter 19 Local, Anytime Steps of Controlled Evolution

Discovering Entities

Keep Grounded

Notice Unnamed Data

Fixing Identifiers

Table of Contents

Underidentification

Overidentification

General Misidentification

Extraneous Identifiers

Seeking Descriptors

Asking the Question

Possible Answers to the Question

Giving the Users Extra Help

Promoting Attributes

Promoting a Plural Attribute

Promoting a Singular Attribute

Relocating Misplaced Descriptors

Factors for Characterizing Descriptor Misplacement

Ways to Detect Descriptor Misplacement

Exercises

Chapter 20 Global, Anytime Steps of Controlled Evolution

Redrawing the Diagram

Altering the Overall Style of an LDS

Changing the Level of Abstraction

Combining Multiple Short Paths

Collapsing a Taxonomy

Collapsing Subordinates Out

Guidelines for Increasing or Decreasing Abstractness

When You Can and Cannot Use Abstract Shapes

Exercises

Chapter 21 Conversations About Dairy Farming

Meeting with Users from the General Offices

Meeting with Veterinary Epidemiologists

Table of Contents

Meeting with Economic Analysts

Exercises

Story Interlude

Chapter 22 Constraints

Constraint Definition Requires a Stabilized Data Model

Many Candidate Constraints Turn Out to Be False

Many Constraints Subject a Data Model to Premature Obsolescence

Worthy Constraints

Constraints and Shifting the Burden

Summary and Final Thoughts

Exercise

Chapter 23 LDS for LDS

The Meta-LDS

Entities and Attributes

Entities Have Descriptors

Relationships and Links

Identifiers

Discussion

Anchoring Your Understanding with Instances

Anticipating How the Meta-LDS Might Evolve

Constraints on the Meta-LDS

Summary

Exercises

Chapter 24 Decisions: Designing a Data-Modeling Notation

Overall Decisions

What Is the Purpose?

What Concepts Are Modeled?

Table of Contents

What Are the Names of the Modeled Concepts?

Should a Model Include Behavior?

What Graphical Notations Should Be Used?

Decisions About Entities

Should an Entity Name Characterize One Instance or Many?

Should There Be Different Notations for Different Kinds of Entity?

Should Each Entity Have an Identifier?

Decisions About Identifiers

Should All Identifiers Be Arbitrary?

How Should Identifiers Be Annotated?

Can Identifiers Include Links?

Can an Entity Have Multiple Identifiers?

Must an Entity Have Multiple Identifiers?

How Are Multiple Identifiers Annotated?

Decisions About Attributes

Is There a Difference Between Entities and Attributes?

Do Attributes Belong on the Diagram?

Do Data Types Belong on the Diagram?

Do Scales Belong on the Diagram?

Are Plural Attributes Allowed?

Are Foreign Key Attributes on the Model?

Are Type-Level Attributes Allowed?

Decisions About Relationships

Are All Relationships Binary?

Is There a Difference Between Relationships and Entities?

Can a Relationship Have Attributes?

Can a Relationship Have Links?

Can a Relationship Have an ID? Must a Relationship Have an ID?

Is There a Difference Between Relationships and Entities? Revisited

Table of Contents

Do Relationships Have Names?

Should There Be Different Notations for Different Kinds of Relationships?

Should Relationship Names Be Verbs?

What Does a Relationship Look Like on the Diagram?

Decisions About Links

Do Links Have Names?

Are Link Names on the Diagram?

How Should a Link Label Be Annotated?

Which Links Get Labeled?

What Does Maximum Degree Look Like?

Decisions About Descriptors

Should Entities and Descriptors Share a Namespace?

Decisions About Constraints

Should the Diagram Capture Minimum Degree?

Should the Diagram Capture Intrainstance, Intraattribute Constraints?

Should the Diagram Capture Intrainstance, Interattribute Constraints?

Should the Diagram Capture Intrainstance, Intralink Constraints?

Should the Diagram Capture Intrainstance, Interdescriptor Constraints?

Should the Diagram Capture Triangle Relationships?

Should the Diagram Capture Interinstance, Intradescrptor Constraints?

Should the Diagram Capture Interinstance, Interdescriptor, Intraentity Constraints?

Should the Diagram Capture Interinstance, Interentity Constraints?

Summary and Final Thoughts

Exercises

Chapter 25 LDS and the Relational Model

Relational Databases

Mapping an LDS to a Relational Schema

LDS and Normal Forms

Table of Contents

First Normal Form

Second Normal Form

Third Normal Form

Fourth Normal Form

Fifth Normal Form

Summary

Exercises

Chapter 26 Cookbook: Recipes for Data Modelers

Set Recipes

Graph Recipes

Matrix Recipes

Taxonomy and Near-Taxonomy Recipes

Exercises

Story Interlude

Appendix: Exercises for Mastery

Index