McGraw-Hill Series on Data Warehousing and Data Management

W. H. Inmon
John A. Zachman
Jonathan G. Geiger

Data Stores Data Warehousing AND THE Zachman Framework

MANAGING ENTERPRISE KNOWLEDGE

数据存储、数据仓库和 Zachman 框架

McGraw-Hill Book Co
世界图书出版公司

DATA STORES, DATA WAREHOUSING AND THE ZACHMAN FRAMEWORK

McGraw-Hill Book Co

书　　名：Data Stores, Data Warehousing, and the Zachman Framework
作　　者：W. H. Inmon, J. A. Zachman, et al.
中 译 名：数据存储、数据仓库和 Zachman 框架
出 版 者：世界图书出版公司北京公司
印 刷 者：北京中西印刷厂
发　　行：世界图书出版公司北京公司 (北京市朝阳门内大街 137 号　100010)
开　　本：大 32 开　850 × 1168　　印　张：12
版　　次：1999 年 4 月第 1 版　　1999 年 4 月第 1 次印刷
书　　号：7-5062-4121-8/ TP•42
版权登记：图字 01-99-0014
定　　价：58.00 元

世界图书出版公司北京公司已获得 McGraw-Hill Book Co. Singapore 授权在中国境内独家重印发行。

Library of Congress Cataloging-in-Publication Data

Inmon, William H.
Data stores, data warehousing, and the Zachman Framework : managing enterprise knowledge / W. H. Inmon, John A. Zachman, Jonathan G. Geiger.
p. cm.
Includes biographical references and index.
ISBN 0-07-031429-2
1. Database management. 2. Management information systems. I. Zachman, John A. II. Geiger, Jonathan G. III. Title.
QA76.9.D315376 1997
005.74'068—dc21 97-1276
CIP

McGraw-Hill
A Division of The McGraw-Hill Companies

Copyright © 1997 by The McGraw-Hill Companies, Inc. All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.

3 4 5 6 7 8 9 0 FGR/FGR 9 9 0 2 1 0 9 8 7

ISBN 0-07-031429-2

Copyright © 1999 by McGraw-Hill Companies, Inc. All Rights reserved. Jointly Published by Beijing World Publishing Corporation/McGraw-Hill. This edition may be sold in the People's Republic of China only. This book cannot be re-exported and is not for sale outside the People's Republic of China.

IE ISBN: 0-07-116386-7

McGraw-Hill books are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please write to the Director of Special Sales, McGraw-Hill, 11 West 19th Street, New York, NY 10011. Or contact your local bookstore.

Information contained in this work has been obtained by The McGraw-Hill Companies, Inc. ("McGraw-Hill") from sources believed to be reliable. However, neither McGraw-Hill nor its authors guarantee the accuracy or completeness of any information published herein, and neither McGraw-Hill nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw-Hill and its authors are supplying information, but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

FOREWORD

John Zachman is the foremost expert on comprehensive architectures for Information Technology. His profound work, culminating in the Zachman Framework, has been the cornerstone for many IT departments struggling with the difficult problems of building, integrating, and maintaining their complex systems.

For years, John spoke to individual companies, in seminars and conferences the world over, educating and entertaining us with his brilliance. And for years, people have begged him to write "the book" on the Zachman Framework. I am so very pleased to say that—at long last—the book has indeed been written.

The integration of John's framework with Bill Inmon's important work in data warehousing and operational data stores is an unbeatable combination. Their ideas mesh perfectly giving all of us the ideal roadmap to build our operational systems and strategic decision support systems. Now we have the complete architecture from which to build sustainable systems.

Finally, it is the practical experiences from Jonathan Geiger that provide us with the real-world advice on implementing the Framework. His pragmatic additions to the book reflect the many data warehouses and operational systems he has built using the Zachman Framework and Inmon methodology. His recommendations are of the highest quality.

I hope that you find this book to be as useful as I do. It is, without a doubt, one of the most important books published for IT professionals everywhere. Thank you, John, Bill, and Jonathan, for providing us with such a valuable addition to our IT libraries!

—Claudia Imhoff
President and Founder
Intelligent Solutions, Inc.

此为试读,需要完整PDF请访问: www.ertongbook.com

PREFACE

Data . . . Information . . . Knowledge. There is a difference.

As companies prepare to move into the twenty-first century, data and information alone will not suffice. Knowledge will be one of the distinguishing assets of successful corporations. Data simply reflects details found throughout a company, and information provides meaning to the data. Knowledge entails the synthesis of information to provide a corporate entity with an improved awareness and understanding of itself, its human and other resources, and its business environment. Armed with this wisdom, companies will be better prepared to compete in an ever-changing global economy.

The data warehouse, operational data store, and client/server systems have been instrumental in helping companies obtain, manage, and leverage data, transforming it into valuable information. These concepts have been made possible through the integration of advances in database management systems, programming and data access languages, computer hardware and software, and communications facilities. These technologies, by themselves, do not yield the right information . . . at the right time . . . at the right place . . . in the right form. To achieve that end, to consistently convert information into knowledge, companies need to adopt an architected approach.

The state of computer utilization evolved with changes in technology. Companies adopted more powerful computers, computer terminals, client/server tools, database management systems, and the like, to improve productivity and to process data into information. With this technology-driven approach, companies moved through three generations of an evolving computer environment. The Formation Generation introduced the computer to the commercial sector. Recognizing its value, companies expended great sums of money on additional computers and support personnel and facilities, leading to a Proliferation Generation. Money alone did not overcome the unsatiable demand for computer applications. As the backlog grew, business units took matters into their own hands and built their own systems. The Dispersion Generation is typified by widespread use of computers throughout the company without a focused effort directed toward managing them.

Clearly, a better approach is needed. That approach, the Unification

Generation, requires a shift in management emphasis. Companies must shift from managing the computer and the associated technologies to managing the environment within which the computers operate and to managing data as a corporate resource. With this shift, the vast amount of bits and bytes contained in a company's computers can be transformed from data into information, and from information into knowledge.

The Zachman Framework provides the foundation for architecting such an environment. Developed by observing work performed in other disciplines, the Framework addresses all aspects of the Enterprise as a whole and of the Enterprise as viewed by the information technology organization. An architecture such as the one provided by the Framework is required to satisfy modern needs. This architecture enables companies to evolve from their legacy system environment into one which employs data warehouses, operational data stores, and client/server technology to meet their changing needs in a competitive environment.

This book is divided into three Parts. In Part One, the importance of architecture is stressed, and the evolution of the computer environment is traced through the four generations. This review of history points out why some of the approaches of the past have failed, and how an architected environment can help.

Part Two describes the Zachman Framework and demonstrates its role in transforming the traditional *data administration* and *information resource management* role into a *knowledge management* role. The transformation is illustrated by describing the application of the Framework as a generic thinking tool for improving understanding and for making better decisions about designing and managing change for any complex object. In this context, the complex object may be a product, the enterprise itself, the architecture of the enterprise, or the knowledge base of the enterprise.

After describing the Zachman Framework and its transformation role, the Framework is applied directly to develop a strategy and methodology for building and managing data warehouses. Within these chapters, concrete examples of guiding principles and model components are developed. The foundation laid through the Framework enables companies to extend their data warehouse as a means for building an architected knowledge management environment.

Part Three applies the first parts. First, the Framework-based methodology is used to develop a data warehouse. The value of the architected approach is then demonstrated as the complexity of the data warehouse increases, and as the information gained from its development is used to

build an operational data store. The Framework is then used to help define the metadata needs. Chapter 11 portrays the data warehouse and operational data as a means to an end. Using the Framework, these serve as a foundation to help companies migrate to an architected environment which supports processes and systems that are engineered to meet the challenges posed by the dynamic business environment.

The book is designed primarily for managers who are interested in improving the management of complex technological environments. It emphasizes the criticality of an architecture-based approach and provides a conceptual description of the Zachman Framework as the basis for gaining greater business value from information systems and electronic databases. Parts Two and Three are also aimed at the practitioner. After developing a Framework-based strategy and methodology in Chaps. 5 and 6, respectively, the value of the Framework is applied toward building a data warehouse and an operational data store, migrating systems to a new environment, and supporting reengineering efforts and total quality management. This book explains key concepts and their importance and provides information to assist practitioners in understanding the critical steps in the process.

There are many people who have contributed directly and indirectly to this effort. The authors wish to acknowledge particularly the following individuals for their contributions:

- Jim Ashbrook, Prism Solutions
- Arnie Barnett, Barnett Data Systems
- Genevieve Blair
- Jeff Colen, Club Runner
- Larry English, Information Impact International
- Lowell Fryman, Intelligent Solutions
- Tom Hokel, Framework Software
- Sam Holcman, Pinnacle Systems
- David Houssian, Utility Partners
- Claudia Imhoff, Intelligent Solutions
- John Ladley, META Group
- Sue Osterfelt, NationsBank
- Ron Ross, *Database Newsletter*
- Phil Seeger, Florida Power & Light
- Ron Shelby, American Express

- Ben Stern, Israel Electric
- Bill Wallace, Cambridge Technologies
- J. D. Welch, DataWing Consulting Services
- Ed Young, D2K

The number of companies which are using the Zachman Framework is continuously increasing. In 1995, the Zachman Institute for Framework Advancement (ZIFA) in Pinkney, Michigan was formed to provide a forum for companies to build on each others' successes in applying the Framework. Over 100 people attended the first ZIFA conference in June 1996. The energy level of the participants was exhilarating. The message was clear: Architecture is not an option—it is a mandate for companies which want to survive and thrive in the twenty-first century. Additional information about the Framework and its application can be obtained by contacting the Institute directly through the Internet at www.zifa.com.

—W. H. INMON
JOHN A. ZACHMAN
JONATHAN G. GEIGER

CONTENTS

Foreword xiii
Preface xv

PART ONE THE NEED FOR AN ARCHITECTURE

Chapter 1 A Look at History with a View to the Future 3

Four Generations 4
Formation 5
Computer Technology 6
Data Management 7
Programming Languages 7
Methodology 7
Assessment 8
Proliferation 10
Computer Technology 10
Data Management 11
Programming Languages 12
Methodology 13
Assessment 15
Dispersion 18
Computer Technology 19
Data and Object Management 19
Programming Languages 21
Methodolgy 22
Assessment 23

Chapter 2 Satisfying Modern Information Needs 29

Need for an Architecture 30
Information Systems Architecture 31
Zachman Framework 39
Migration of Legacy Systems 39
Influencing Factors 40
Role of Technology 41

Unification 42
Computer Technology 43
Data and Object Management 44
Programming Languages 47
Methodology 48
Assessment 48

PART TWO ARCHITECTURAL FRAMEWORK

Chapter 3 The Zachman Framework 55

Zachman Framework Overview 56
Perspectives 57
Dimensions 57
Perspectives—Rows of the Zachman Framework 60
Planner's Perspective 61
Owner's Perspective 63
Designer's Perspective 64
Builder's Perspective 65
Subcontractor's Perspective 66
Dimensions—Columns of the Zachman Framework 71
Entities (What?) 71
Activities (How?) 72
Locations (Where?) 73
People (Who?) 74
Times (When?) 75
Motivations (Why?) 76
Rules of the Zachman Framework 78
Dimension Importance 78
Dimension Simplicity 80
Dimension Uniqueness 81
Perspective Uniqueness 82
Cell Uniqueness 82
Dimension Necessity 85
Logic Recursiveness 87

Chapter 4 A Framework for Managing Enterprise Knowledge 93

Observations about the Framework 94
The Product Framework 96

此为试读,需要完整PDF请访问: www.ertongbook.com

The Enterprise Framework 98
Relationship to Product Framework 99
Application 101
Engineering Administration 103
Data Administration 104
Data Warehouse and Operational Data Store 105
The Challenge 106
Information Systems Department Framework 107
The Problem 108
Application 111
Repository Framework 116
Application 117
Analogy to Engineering Administration 121

Chapter 5 Establishing the Data Warehouse Strategy 127

What Is a Guiding Principle? 128
Motivations Dimension 130
Funding 131
Service Levels 131
Times Dimension 132
Historical View 132
Historical Information 134
Data Currency 134
Locations Dimension 135
Servers 135
Indirect Data Access 137
Activities Dimension 139
Data Warehouse Methodology 139
Application Development Methodology 141
People Dimension 142
Direct Data Access 142
Interface 143
Security 143
Entities Dimension 143
Quality 144
Data Sources 145
Transformation and Integration 145
Data Orientation 145
Metadata 146

Chapter 6 Zachman Framework–Based Methodology 149

Methodology Overview 150
Planner's View 152
Commitment to Proceed 152
Owner's View 157
Data Model Analysis 158
Designer's View 164
Breadbox Analysis 164
Technical Assessment 166
Subject Area Analysis 168
Builder's View 169
Source System Analysis 169
Technical Environment Preparation 172
Data Warehouse Design 173
Programming Specifications 176
Technical Environment Preparation 178
Subcontractor's View 179
Programming 179
Technical Environment Preparation 179
Builder's View Revisited 179
Population 179
Metadata 180
Tool Support 180

PART THREE APPLICATIONS

Chapter 7 Building the First Data Warehouse 185

The Development Life Cycle 186
Commitment to Proceed 187
Motivations 188
People and Locations 189
Entities 191
Activities and Times 191
Risk Assessment 192
Project Sponsor Implications 195
Project Manager Implications 195
Source System Analysis 195
Data Model Analysis 196
Removing Operational Data 196

	Identifying Business Entities and Attributes	197
	Adding a Time Dimenstion	198
	Risk Assessment	198
	Technical Assessment and Technical Environment Preparation	199
	Breadbox Analysis	200
	Subject Area Analysis	200
	Source System Analysis	200
	Data Warehouse Design	200
	Specifications	202
	Activities	202
	Motivations	202
	People	203
	Locations	204
	Times	204
	Technical Model Culmination	205
	Programming	205
	Entities	206
	Activities	207
	Locations	207
	People	207
	Times	207
	Motivations	208
	Out-of-Context Model Culmination	208
	Technical Model Revisited—Population	209
Chapter 8	Extending the Data Warehouse	213
	Commitment to Proceed	214
	Motivations	214
	People	216
	Data Model Analysis	217
	Entities	218
	Attributes	219
	Business Rules	219
	Data Steward	220
	Subject Area Analysis	222
	Source System Analysis	222
	Data Warehouse Design	225
	Staging Area	226
	Star Schema	226
	Additional Subject Areas	228

Chapter 9 Building the Operational Data Store 233

Contrasts and Similarities 234
Activities 235
Planner's Perspective 236
Owner's Perspective 238
Designer's Perspective 239
Builder's Perspective 243
Subcontractor's Perspective 243
Motivations 243
People 243
Entities 246
Business Model 246
System Model 247
Technical Model 247
Out-of-Context Model 248
Times 248
Locations 249
Operating Environment 250

Chapter 10 Metadata 253

Importance of Metadata 254
Metadata about the Data Warehouse 255
Metadata about People 257
Metadata about Entities 258
Metadata about Activities 260
Metadata about Locations 261
Metadata about Motivations 262
Metadata about Times 263
Combined View 264
Metadata Capture and Maintenance 265
Metadata for the Initial Data Warehouse 266
Metadata for Multiple Data Warehouses 267

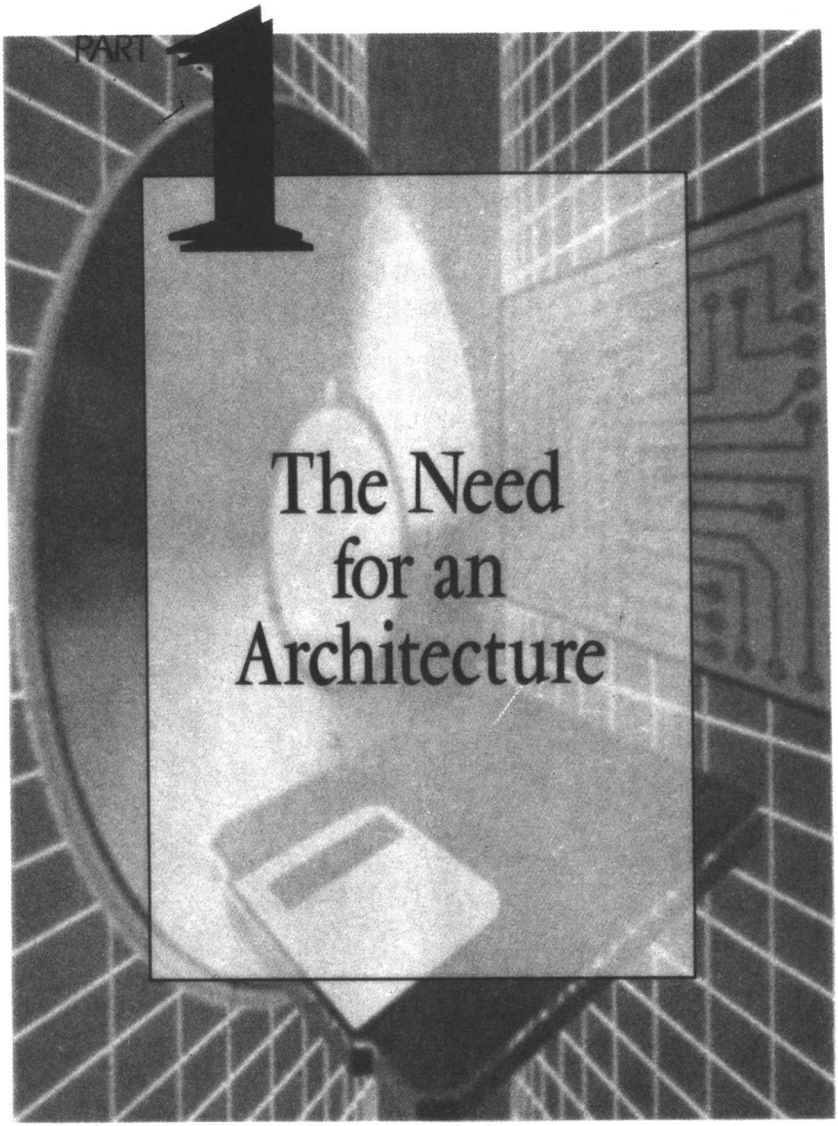
Chapter 11 The Data Warehouse as a Means to an End 269

Migrating Legacy Systems 270
The Problem 270
The Opportunity 270
Scope Definition 271

Business Model 274
Technical Model 275
System Model 275
Technical Model Revisited 279
Out-of-Context Model 279
Impact 280
Summary of System Migration 281
Helping Companies Deal with a Dynamic Business Environment 281
Support from the Zachman Framework 283
Support from the Data Warehouse 286
Support from the Operational Data Store 287
Support from New Operational Systems 287
Total Quality Management 288
Management of Information and Data 290
Competitive Comparisons and Benchmarking 291
Analysis and Use of Company-Level Data 291
Management by Fact 292
Root Cause Analysis 292

Chapter 12 Knowledge 295

Glossary 297
Bibliography 329
Index 345



此为试读,需要完整PDF请访问: www.ertongbook.com