

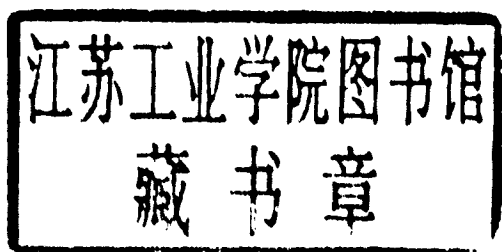
Analysis, Design, & Implementation of Data Dictionaries



Ken S. Brathwaite

Analysis, Design, and Implementation of Data Dictionaries

Ken S. Brathwaite



McGraw-Hill Book Company

New York St. Louis San Francisco Auckland
Bogotá Hamburg London Madrid Mexico
Milan Montreal New Delhi Panama
Paris São Paulo Singapore
Sydney Tokyo Toronto

Library of Congress Cataloging-in-Publication Data

Brathwaite, Ken S.

Analysis, design, and implementation of data dictionaries / Ken S. Brathwaite.

p. cm.

Bibliography: p.

Includes index.

ISBN 0-07-007248-5 : \$39.95

1. Data base management. 2. Data dictionaries. I. Title.

QA76.9.D3B688 1988

005.74'2—dc19

87-20983
CIP

Copyright © 1988 by McGraw-Hill, 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.

1234567890 DOC/DOC 893210987

ISBN 0-07-007248-5

The editors for this book were Theron Shreve and Georgia Kornbluth, the designer was Naomi Auerbach, and the production supervisor was Dianne Walber. It was set in Century Schoolbook by BYRD Press.

Printed and bound by R. R. Donnelley & Sons Company.

Analysis, Design, and Implementation of Data Dictionaries

*To Monique, Marguerite,
Michele, and Melanie.
I love you all.*

To My Readers

The express purpose of this book is to discuss in a meaningful way the analysis, design, and implementation of data dictionaries.

The book starts off with a discussion of the environments in which data dictionaries are used. Some of these include environments in which data analysis is done in a methodical fashion and extensive use is made of the data dictionary to record the results of the analysis phase. Another is the information center, where users can obtain solutions for their problems. The data dictionary is a valuable tool in defining and describing the essential data elements that may be of interest to the user—hence the discussion of these environments.

The intent of this book is to enable readers to develop their own in-house data dictionaries to meet their own unique requirements. Readers will quickly determine that, in some chapters, more emphasis is placed on the discussion of background material than on the actual development of the data dictionary for the particular environment. The reasoning behind this approach is that if the reader understands the environment, then the development of the dictionary for that environment becomes a simple matter. Thus the chapters on the use of data dictionaries in online and distributed data-processing, office automation, and data-security environments place heavy emphasis on the understanding of these environments.

The design aspect of data dictionaries centers on the metadata (entries) and standards that must be incorporated into average data dictionaries. Several chapters, including Chapter 3, are devoted to these entries and standards. Readers may want to incorporate all these entries into their data dictionaries or may prefer to select those entries that best describe their data requirements.

A data dictionary can be implemented by creating a database from the entries described in this book, using any of the currently available utilities to load the database and a teleprocessing (TP) monitor to retrieve information from the data dictionary. Chapter 4 discusses a

typical implementation of a data dictionary and gives an example of the entries retrieved from the dictionary.

The book ends with case histories of actual usage of data dictionaries in two corporations. The case histories are unedited and were obtained by sending surveys to various companies.

PREFACE

The work reported in this book was developed from notes I used in teaching a course for graduate students at the University of Alberta and from research conducted at Alberta Government Telephones since 1981. The book is intended to serve as a practical guide for workers at all levels who are responsible for designing and implementing data dictionaries.

The main objective of this book is to provide material that is essential for the efficient design and implementation of data dictionary systems. The book provides workers with the tools necessary for either selecting data dictionary designs that will be adequate for their environments or developing their own in-house dictionary systems.

The content of the book has been considerably enhanced by my experience as a member of the American National Standards Institute (ANSI) X3H4 Committee, responsible for establishing standards for data dictionaries.

The report on the usage of data dictionaries in various companies reflects the results of surveys sent to the Fortune 500 companies in the United States and the Financial Post 500 companies in Canada. The surveys sought to determine the type of dictionary in use (whether a manufacturer's or one developed in-house) in each company, the history and usage of the dictionary, whether the dictionary was used in an active or a passive manner, the maintenance requirements of the dictionary, and the special reports produced by the dictionary.

Part I introduces the environments in which a dictionary may be used. Chapter 1 details those environments and covers such topics as (1) data analysis and functional analysis, (2) logical and physical design, (3) documentation, (4) data security and integrity, and (5) user and information center requirements.

Chapter 2 introduces the basic concepts of the data dictionary and covers the definition of a data dictionary, active versus passive data

dictionaries, case histories of data dictionary uses, entries and contents of data dictionaries, and the concept of metadata.

Chapter 3 deals with the design of in-house dictionaries. It discusses the reasons for developing your own dictionary, the planning requirements for development of the dictionary, and an implementable structure for an in-house dictionary. Chapter 4 discusses a method of entering data into an in-house data dictionary.

Chapters 5 and 6 discuss the results of surveys conducted in two environments—distributed data processing and office automation—to determine how dictionaries were used in those environments. Chapter 7 deals with the security aspects of the data dictionary. It illustrates not only how to secure the contents of the data dictionary but how to use the data dictionary to achieve both data security and physical security.

Chapter 8 discusses the dictionary as a tool for automated physical database design. It shows how a dictionary can be used to store information about the attributes, entities, and logical schemas which are used to build the physical databases. The designer can use the known relationships among these entries to automatically group and produce the building blocks required for database design. Chapter 9 shows how standards are developed and used in a data dictionary environment.

Chapter 10 discusses performance indicators in data dictionaries. It shows ways in which users can collect and use statistics to determine how well a dictionary is performing in relation to the needs of the particular environment.

Part II presents case histories of dictionaries and how they are being used in various organizations.

ACKNOWLEDGMENTS

I am grateful for the comments and suggestions I received from Vic Howard, Stan Locke, and Francis Chin. The initial draft of this manuscript was ably typed by Jane Cuffy. Her efforts are appreciated.

Ken S. Brathwaite

Analysis, Design, and Implementation of Data Dictionaries

To My Readers xlii
Preface xv

1.30	Staffing the Information Center	25
1.31	Problems Affecting the Information Center	25
1.32	Data Administration and Database Administration	26
1.33	Establishing the Data-Administration Function	26
1.34	The Functions of Data Administration	27
1.35	The Activities of Data Administration	27
1.36	Functions of Database Administration	28
1.37	Standards and Naming Conventions	28
1.38	Change Control	29
1.39	Hardware and Software Performance Statistics	30
1.40	Consistency Checks	30
1.41	Summary	31
2.	Concepts of Data Dictionaries	33
	Introduction	33
2.1	What Is a Data Dictionary?	33
2.2	The Concept of Metadata	33
2.3	Active versus Passive Data Dictionaries	34
2.4	Case Histories of Data Dictionary Use	34
2.5	Data Dictionary Use at Company A	35
2.6	An Example of Management Uses of the Data Dictionary	35
2.7	Company A's Approach to Using the Data Dictionary	35
2.8	Company A's Use of the Data Dictionary for Business Planning	36
2.9	Data Dictionary Use at Company B	36
2.10	Company B's Use of the Extensibility Feature of the Data Dictionary	37
2.11	The Metadata Entries for a Typical Dictionary	37
2.12	The Data Dictionary as a Directory	40
2.13	Maintenance of the Data Dictionary	40
2.14	Access Authorization of Users	41
2.15	Consistency and Validation Checks	41
2.16	Directory and Data-Storage Features	41
2.17	Relationship or Membership of an Entry	42
2.18	Environmental Data	42
2.19	Active and Passive Data Dictionaries	42
2.20	Design of Data Dictionaries	43
2.21	Control and Audit Features of Data Dictionaries	43
2.22	Data Dictionary Standards	44
2.23	The Data Dictionary as a Tool for Data Analysis	45
3.	Design of In-House Data Dictionaries	47
	Introduction	47
3.1	The Entity-Relationship Approach	47
3.2	Introduction to the ER Data Dictionary	48
3.3	Definitions and Terms Used in the Data Dictionary	50
3.4	Structure of the Data Dictionary	51
3.5	The Attribute Definition	51
3.6	Application of the Data Dictionary	56
3.7	Examples of Use of an In-House Data Dictionary	57
3.8	Retrieval from the Data Dictionary	57
3.9	Practical Uses of In-House Data Dictionaries	59
3.10	Summary	60

4. A Method for Entering Data into the Data Dictionary	63
Introduction	63
4.1 Data Entry	63
4.2 Entering Section Headings	64
4.3 Entering Text Data	64
4.4 Entering Paragraphs	64
4.5 Entering Tables or Lists	65
4.6 Entering Combinations of Paragraphs and Tables	65
4.7 Editing Members of a PDS	65
4.8 Editing an Existing Member	66
4.9 Creating a New Member	66
4.10 Deleting a Member	67
4.11 Browsing a PDS	67
4.12 Selecting a Member from a PDS	67
4.13 Printing Members from a PDS	68
4.14 Examples of a Created Member of a PDS	68
4.15 Summary	86
 5. Use of Data Dictionaries in Online and Distributed-Data-Processing Environments	 87
Introduction	87
5.1 Unauthorized Access	88
5.2 Loss of Audit Trail	89
5.3 Transmission Errors and Controls	90
5.4 Data-File Controls	90
5.5 Output Controls	91
5.6 Evaluation of Controls	91
5.7 The Distributed Data-Processing Environment	92
5.8 Replication and Partition in a DDP Environment	93
5.9 Data Dictionary and Directories for DDP	94
5.10 Using a Schema to Describe DDP Data	95
5.11 Dynamic Distribution of Data	96
5.12 Transaction Management	97
5.13 The Role of the Data Dictionary in Defining the DDP System	98
5.14 Data Dictionary Definitions of DDP Data	99
5.15 Data to Be Captured for Each Category	99
5.16 Summary	100
 6. Data Dictionaries in Office Automation Environments	 101
Introduction	101
6.1 Technology Used in Electronic Office Workstations	102
6.2 Improvements in Office Personnel Support	103
6.3 Network Architecture for Office Automation	103
6.4 The Open Systems Interconnection Model	104
6.5 Local Area Networks	104
6.6 The Private Branch Exchange	104
6.7 Functional Areas in Office Automation	105
6.8 The Data Dictionary's Role in Storage and Retrieval	106

7. Security and Integrity of Data Dictionaries	107
Introduction	107
7.1 Definitions of Security, Privacy, and Integrity	107
7.2 Conducting a Threat Analysis	109
7.3 Examples of Survey Questions	109
7.4 Protection Mechanisms	110
7.5 Access Management	110
7.6 Privacy Transformations	110
7.7 Cryptographic Controls	111
7.8 Security Kernels	111
7.9 Access Matrix	111
7.10 Protecting the Data Dictionary	113
7.11 The Data Dictionary as a Protection Mechanism	113
8. Database Design Using the Data Dictionary	115
Introduction	115
8.1 Review of Existing Methodologies	115
8.2 Detailed Discussion of Database Design	116
8.3 Inputs to the Design Process	117
8.4 The Entity-Relationship Methodology	118
8.5 The Data-Analysis Phase	119
8.6 Conducting the Data-Analysis Phase	120
8.7 The Entity Model	120
8.8 Role of the Data dictionary in Data Analysis	121
8.9 Role of the Data Dictionary in Entity Modeling	121
8.10 Logical Schema Design	121
8.11 Role of the Data Dictionary in Physical Design	122
8.12 The Data Dictionary as a Directory	122
9. Developing Data Dictionary Standards	123
Introduction	123
9.1 Categories of Data Dictionary Standards	123
9.2 Standard Formats for Data Dictionary Entries	124
9.3 Standards for Programs Interfacing with a Data Dictionary	124
9.4 Security Standards	125
10. Performance Indicators of Data Dictionaries	127
Introduction	127
10.1 Types of Reports Produced	127
10.2 Ability to Produce Ad Hoc Reports	128
10.3 Ease of Entering and Retrieving Data	128
10.4 Ability to Check Accuracy of Contents	128
10.5 Ability to Associate Attributes with Entities	128
10.6 Interaction of Processes and Data Elements	128
10.7 Data Dictionary Summary Reports	128
PART 2 Case Histories of Data Dictionary Usage	131
Case History 1	133
Case History 2	155

Appendix A	Entity Class Definition Contents	161
Appendix B	Logical Group Definition Contents	169
Appendix C	IMS Database Definition Contents	183
Appendix D	Logical Schema Definition Contents	193
Bibliography		205
 Index		 207

Part

1

Concepts of Data Dictionaries

