# 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



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

# CONTENTS

Part

# 1

# Concepts of Data Dictionaries