Developing a Data Warehouse for the Healthcare Enterprise

Lessons from the Trenches

Third Edition

Bryan Bergeron, MD, Editor

Hamad Al-Daig, MBA
Osama Alswailem, MD, MA
Enam UL Hoque, MBA, PMP, CPHIMS
Fadwa Saad AlBawardi, MS

HIMSS

CRC Press
Taylor & Francis Group
A PRODUCTIVITY PRESS BOOK

Developing a Data Warehouse for the Healthcare Enterprise

Lessons from the Trenches

Third Edition

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Developing a Data Warehouse for the Healthcare Enterprise

Lessons from the Trenches

Third Edition

Bryan Bergeron, MD, Editor

Hamad Al-Daig, MBA

Osama Alswailem, MD, MA

Enam UL Hoque, MBA, PMP, CPHIMS

Fadwa Saad AlBawardi, MS

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A PRODUCTIVITY PRESS BOOK

HIMSS

Taylor & Francis
Boca Raton London New York
A CRC Press title, part of the Taylor & Francis imprint, a member of the Taylor & Francis Group, the academic division of T&F Informa plc

Published in 2018 by CRC Press
a Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2018 by Healthcare Information and Management Systems Society (HIMSS).
CRC Press is an imprint of Taylor & Francis Group

No claim to original U.S. Government works

Printed in the United States of America on acid-free paper 10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-138-50296-3 (Hardback)
International Standard Book Number-13: 978-1-138-50295-6 (Paperback)
International Standard Book Number-13: 978-1-315-14517-4 (eBook)

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

No part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Names: Bergeron, Bryan P., author. | Alswailem, Osama, author. | Al-Daig, Hamad, author. | Hoque, Enam UL, author. | AlBawardi, Fadwa Saad, author.
Title: Developing a data warehouse for the healthcare enterprise : lessons from the trenches / Bryan Bergeron, Osama Alswailem, Hamad Al-Daig, Enam UL Hoque, Fadwa Saad AlBawardi.
Description: Third edition. | Boca Raton : Taylor & Francis, 2018. | Includes bibliographical references and index.
Identifiers: LCCN 2017056424| ISBN 9781138502963 (hardback : alk. paper) | ISBN 9781138502956 (paperback : alk. paper) | ISBN 9781315145174 (ebook)
Subjects: LCSH: Medical care--Information technology. | Medical care--Data processing.
Classification: LCC R858 .B4714 2018 | DDC 610.285--dc23
LC record available at https://lccn.loc.gov/2017056424

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

HIMSS Mission

To lead healthcare transformation through the effective use of health information technology.

© 2013 by Healthcare Information and Management Systems Society (HIMSS).

All rights reserved. No part of this publication may be reproduced, adapted, translated, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

Printed in the U.S.A. 5 4 3 2 1

Requests for permission to make copies of any part of this work should be sent to:
Permissions Editor
HIMSS
33 W. Monroe St., #1700
Chicago, IL 60603-5616
nvitucci@himss.org

The inclusion of an organization name, product, or service in this publication should not be considered as an endorsement of such organization, product, or service, nor is the failure to include an organization name, product, or service to be construed as disapproval.

For more information about HIMSS, please visit www.himss.org.

Contents

Preface

This is the third edition of *Developing a Data Warehouse for the Healthcare Enterprise: Lessons from the Trenches*, the first edition having received the 2008 HIMSS Book of the Year Award. The primary goal of this book is to provide an up-to-date, straightforward view of a clinical data warehouse project at King Faisal Specialist Hospital and Research Centre (KFSH&RC) in Riyadh, Saudi Arabia. Whereas the first two editions emphasized inception and implementation, this third edition looks at the mature project with an eye toward the maintenance phase of the life cycle.

Despite an uptick in data warehouse implementations in the healthcare sector over the past decade, the definitions of exactly what constitutes a data warehouse still vary from one vendor and healthcare enterprise to the next. For the purpose of this book, a data warehouse is defined as a logically central repository for selected clinical and nonclinical data from disparate, often loosely integrated systems throughout the healthcare enterprise. In our case, the logically central repository is also physically central.

From a strategic perspective, the data warehouse is an enabling technology that, when properly implemented, can be leveraged to reduce medical errors, promote patient safety, support the development of an enterprise-wide electronic health record (EHR), and support process/work flow redesign. As such, the upside potential for a successful data warehouse implementation is enormous. However, as with any large-scale, expensive, mission-critical IT project, an inferior implementation can spell disaster for not only the IT department staff but for the healthcare enterprise as a whole.

The venue for our discussion, KFSH&RC, is a large, modern, tertiary-care hospital in Saudi Arabia with an environment that parallels leading-edge U.S. hospitals. The clinical departments, surgical wards, operating rooms, bedside monitors, and even the formularies are indistinguishable from those in a tertiary-care hospital in Boston, New York, or San Francisco. There is even

a Starbucks in the main lobby, albeit with palm trees and camels on the souvenir coffee mugs.

More importantly, the IT environment is indistinguishable from the best in the United States, with hardware from the likes of IBM and HP, and an EHR system from Cerner. Given this infrastructure, which includes a data warehouse, it's no surprise that KFSH&RC is the first HIMSS Analytics Stage 7–certified hospital in the Middle East. Moreover, KFSH&RC is leveraging the data warehouse in a proteomics initiative that has brought the enterprise to the forefront of translational medicine.

In addition to reviewing the experiences at KFSH&RC, we examine the value of the data warehouse from the U.S. perspective. We discuss the increasing role of data analytics in supporting an era of increased accountability and personal expense for care in the United States. As a result, the lessons learned should have both domestic and international appeal and applicability.

This book is written for the HIMSS membership—including chief information officers (CIOs), IT managers, and hospital administrators—involved in medical error reduction, patient safety, EHR implementation, and process improvement. It is designed as a road map for healthcare enterprise executives and IT managers contemplating or already involved in a data warehouse implementation. Although the contributors are obviously biased proponents of data warehouse technology, they are quick to point out some of the difficulties and limitations faced during the implementation process and ways to either avoid or overcome them.

The chapters, written by those responsible for different aspects of the project, tell all from personal, hands-on experience. The original contributors have updated their respective chapters to reflect changes since the second edition. The timely update makes this book a must-have for owners of the first and second editions, as well as new readers.

This book is unique in that it provides the perspectives of several key stakeholders in the data warehouse development project at KFSH&RC, from the initial vision to release. We provide the view of the CIO (Hamad Al-Daig), the medical informatician (Osama Alswailem), the technical manager (Enam UL Hoque), the senior program analyst (Fadwa AlBawardi), and the external consultant (Bryan Bergeron). The internal parallels and occasional contradictions exemplify the challenges readers should consider in their own data warehouse development projects. As such, this book also provides insight into the inner workings of a large healthcare enterprise—in itself a valuable resource for healthcare IT professionals.

Developing a Data Warehouse for the Healthcare Enterprise is structured as stand-alone chapters written from different perspectives. Readers are forewarned that, unlike some edited collections that strive for a single voice and perspective, there are numerous points of view that are, on occasion, in apparent contradiction in approach or ranking of importance. These differences in perspective are celebrated and emphasized to illustrate the real-world differences in how a CIO approaches an implementation challenge compared with, for example, a consultant or information systems architect.

Chapter 1, "Here, There Be Monsters," explores the risks and potential upsides of embarking on a data warehouse initiative. It serves as both a sanity check and a gut check for those contemplating the move.

Chapter 2, "The Data Warehouse as Feeder to Data Analytics and Business Intelligence: The Good, the Great, the Bad and the Ugly," explores the relationship of data analytics and business intelligence to decision support and compares decision support based on a data warehouse versus disparate sources.

Chapter 3, "Enterprise Environment," provides an overview of our enterprise environment from an operational prospective, including clinical load, IT infrastructure, and organizational structure.

Chapter 4, "Vendor Selection and Management," provides an overview of request for proposal (RFP) development, vendor selection, and the management processes that were integral to the development of the data warehouse.

Chapter 5, "Development Team," provides an overview of the human resources involved in the data warehouse project, from team formulation to the assignment of roles and responsibilities.

Chapter 6, "Planning," provides an overview of the preparation that went into data warehouse implementation.

Chapter 7, "Design," provides an overview of our technical design, including the data model, logical, and physical architecture; the extraction, transformation, and loading (ETL) process; provision for backup and recovery; and reporting.

Chapter 8, "KPI Selection," explores the process used to determine the most appropriate key performance indicators (KPIs) for our data warehouse implementation.

Chapter 9, "Implementation," describes the highlights of our implementation process, including the ETL build, the online analytical processing (OLAP) build, and user acceptance testing.

Chapter 10, "Post-implementation Organizational Structure," describes the plans defined by management to address the issues of ownership, roles, and responsibilities associated with the data warehouse.

Chapter 11, "Report Life Cycle," defines the data warehouse–based reporting system life cycle, from generation to retirement.

Chapter 12, "Knowledge Transfer," details our approach to managing the transfer of intellectual assets associated with the development of our data warehouse from vendors and consultants to our permanent staff.

The epilogue provides a compilation of the lessons learned from the preceding chapters and discusses their applicability to other data warehouse projects.

Because a data warehouse is a compilation of applications and technologies, numerous acronyms are inevitable. As such, the section entitled "Acronyms" defines the major ones readers are likely to encounter in a data warehouse initiative. Similarly, one of the greatest hurdles for IT executives working with leading-edge technologies in a healthcare organization is using the appropriate terminology when communicating with vendors, engineers, and administrators. The glossary is intended to help bridge the vocabulary gap.

Bryan Bergeron, MD
Boston, Massachusetts

Acknowledgments

Sharing the lessons learned—the hard way—of a major healthcare IT project is no mean feat, even with time for reflection and the benefit of 20/20 hindsight. Actually facing the day-to-day challenges of implementing a multimillion dollar project tests the mettle of even the most seasoned healthcare CIO and management team. Then there are the myriad challenges associated with maintenance, where considerable resources can be spent simply to keep the system functioning even though the world may be in political and economic chaos.

The contributors to this book deserve special acknowledgment for sharing their boots-on-the-ground experiences without the sugar-coating that authors often use for self-promotion. You'll read accounts of what actually transpired during the development of data warehouses—the good, the great, the bad, and the ugly—and take away pearls of wisdom on what to emulate and what to avoid. We believe that knowing what to avoid, and what doesn't work, is at least as important as being able to differentiate what *can* be done from what *should* be done. Learning from the successes and failures of others is more fruitful and less costly than stumbling across your successes and learning from your own mistakes. Moreover, the benefits of such learning accrue to both the individual and the healthcare institution.

Thanks also to those who have been instrumental in the ongoing development of the data warehouse at KFSH&RC, especially Wadood Tafiq, director, Data and Analytics.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

About the Editor

Bryan Bergeron, MD, a fellow of the American College of Medical Informatics, is the author of numerous books, articles, software packages, and patents. He has practiced medical informatics at Massachusetts General Hospital in Boston and taught medical informatics, as well as traditional medical courses, in the Health Sciences and Technology division of Harvard Medical School and MIT for nearly three decades. He has been involved in the development of the data warehouse at King Faisal Specialist Hospital and Research Centre (KFSH&RC) since the inception of the project.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

About the Authors

Hamad Al-Daig, MBA, is a retired chief information officer (CIO) at King Faisal Specialist Hospital and Research Centre (KFSH&RC). He has over 35 years of healthcare IT experience, 28 of those years with KFSH&RC. During his tenure, the hospital attained EMR Adoption Level 6 and ISO 27001 Information Security Management certification. As a leading distinguished professional in healthcare IT, he has contributed to many national-level healthcare IT initiatives, including the development of the national healthcare IT strategy for Saudi Arabia. He is the co-founder and vice president of the Saudi Arabian Health Informatics (SAHI) society. He has been named one of the top 10 CIO innovators of the year by Healthcare Informatics. He is as an emeritus member in good status with KLAS, having served as an international advisory board member for the company. Hamad is CEO of Carelink, a healthcare IT company in Saudi Arabia.

Osama Alswailem, MD, MA, a consultant in family medicine and CIO of KFSH&RC, is the former director of the Medical and Clinical Informatics department. Dr. Alswailem received his medical degree and his board certification in family and community medicine from the College of Medicine, King Saud University, Saudi Arabia. He also obtained a master's degree while completing a postdoctoral fellowship in medical informatics at Columbia University, New York. In addition to his hospital duties, Dr. Alswailem is an assistant professor at Alfaisal University, where he teaches medical informatics.

Enam UL Hoque, MBA, PMP, CPHIMS, is a senior strategic health information consultant at KFSH&RC. He is currently playing an advisory role to the CIO and other C-Suite members within the hospital. Previously, he managed the technical areas of the initial data warehouse development project and introduced performance improvement through the Productivity Analysis

and Benchmarking program for the hospital. He is an IT professional with more than 25 years of experience, holding various positions and managing IT projects of varying sizes within industries ranging from manufacturing and retail to marketing and healthcare in Canada, the United States, and Saudi Arabia.

Fadwa Saad AlBawardi, MS, is the Acting Director, Business and Intelligence Management, ISID, Ministry of National Guard Health Affairs, Riyadh. Ms. AlBawardi formerly worked as a project leader and senior program analyst for the data warehouse section, Healthcare Information Technology Affairs at KFSH&RC in Riyadh. Ms AlBawardi received her MS in computer science at Boston University, Massachusetts, and has been working in the data warehousing/business intelligence areas for several years.

Chapter 1

Here, There Be Monsters

Bryan Bergeron

Contents

Introduction

In medieval Europe, the ocean was the great unknown, fraught with dangers as well as the promise of riches. The life of a captain was a perilous one, in that death could come at any time from bad weather or, even worse, attack by one of the many monsters that supposedly inhabited the ocean. As I've depicted in Figure 1.1, maps of the time often had explicit indications of the dangers lurking in the ocean. Of course, as the centuries passed and new technologies were developed, we came to know the earth as a sphere spinning in space instead of a flat surface, and the sea monsters were either disproven or became sources of food and fuel, fodder for academic studies, and, finally, endangered species on the edge of extinction.

Today, the chief information officer (CIO) of a hospital or healthcare enterprise is a high-risk profession. Based on my ad hoc research, I'd say

Figure 1.1 Common perception of the ocean as depicted on medieval maps.

that two years is the average tenure of a healthcare CIO. This relatively high turnover is due in part to the CIO's inability to deliver on time and on budget, given the ever-changing political, regulatory, and economic foundations of healthcare and the resulting changes in IT projects. There are just too many unknowns—potential monsters out there, if you will—at the time of planning. And when the monsters do appear, it's often too late to change course. Given this reality, why would a healthcare CIO add a data warehouse implementation—recognized as an extremely high-risk proposition[1,2]—to their list of promises? As with the seafaring captains of old, we take the risk because of the promise of significant rewards.

So, what sort of rewards are we talking about? Well, consider that by integrating disparate clinical and administrative data sources into a single source model, a data warehouse can provide clinicians, administrators, and researchers with information from a variety of otherwise noncompatible or poorly integrated sources. Moreover, the data can be had in seconds to minutes, when it has value in clinical decision making, as opposed to days or weeks later. A properly constructed and maintained data warehouse supports *rapid* data mining, *rapid* report generation, and *real-time* decision support—all key components of a full-featured electronic health record (EHR) or electronic medical record (EMR).

Another reward to consider in the risk/reward calculation is recognition by your peers and your institution. Easily at the top of my list for bragging rights is attaining Stage 7 certification of the HIMSS Analytics EMR Adoption Model. As shown in Table 1.1, a data warehouse is one of the four prerequisites for Stage 7 certification. However, there remains the

Table 1.1 Eight Stages of the HIMSS Analytics EMR Adoption Model, with the data warehouse in Stage 7

Stage	*Cumulative Capabilities*
7	Data warehouse Complete EHR CCD (cash concentration and disbursement) transactions to share data Data continuity with emergency department, ambulatory, and outpatient
6	Physician documentation (structured templates) Full clinical decision support system (variance and compliance) Full radiology picture archiving communications system
5	Closed loop medication administration
4	Computerized physician order entry CDS (clinical protocols)
3	Nursing/clinical documentation (flow sheets) CDS system (error checking) PACS available outside radiology
2	Clinical data repository Controlled medical vocabulary CDS May have document imaging Health information exchange capable
1	Lab, radiology, and pharmacy installed
0	No ancillaries installed

economic reality that comes with developing or acquiring a data warehouse; it demands significant human and computer resources over a sustained period of time.

Figure 1.2 provides another view of the data warehouse, in the context of the traditional evolution of healthcare IT capabilities. As shown in the figure, capabilities generally evolve from claims processing to transaction processing to transaction databases, such as those with a particular clinical system. A population of transaction databases—from clinical, administrative, and financial systems—is necessary to feed a comprehensive data warehouse.

As shown in Figure 1.2, there is also an evolutionary path from financial to administrative to clinical computing, culminating in clinical decision support (CDS) capabilities. The latest area of focus at KFSH&RC is IT support

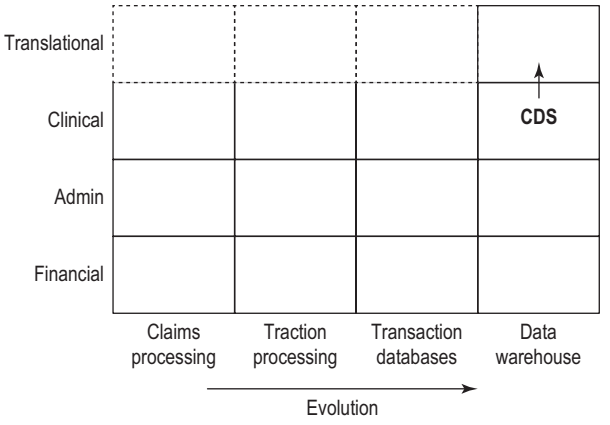

Figure 1.2 Data warehouse in the context of the traditional evolution of healthcare IT capabilities, culminating in CDS capabilities.

for translational medicine, which leverages proteomic data with clinical data residing in the data warehouse.

It's possible to start from scorched earth and purchase an all-in-one EHR solution, complete with a full suite of financial, administrative, and clinical applications, together with a data warehouse, but that's rare. Similarly, it's theoretically possible to skip the financial and administrative processing and go directly to CDS, but such a move is practically impossible. If your institution can't get the bills out, code clinical activities for the maximum legal reimbursement possible, and fulfill administrative obligations, then it won't be operating for long.

For many in the healthcare informatics community, the future of modern medicine is CDS that incorporates both clinical and translational medicine information. For example, when determining whether a patient with a specific genetic predisposition and positive clinical findings should undergo surgical or chemical treatment, the data warehouse serves as an integration enabler for clinical and research computing.

Despite the uncertainties and costs, many in the healthcare industry view adoption of the data warehouse in some form as inevitable once a clear return on investment (ROI) can be identified. In this context, ROI reflects more timely, accurate clinical and administrative data that can form the basis of decisions resulting in cost savings and increased quality of care. These improvements in turn support the goals of *meaningful use*—and are linked to federal incentive funding—as somewhat poorly defined by the HITECH component of the American Recovery and Reinvestment Act of 2009.

One way to address risk, uncertainty, and potential ROI is for organizations involved in data warehouse implementations to share their experiences—whether successful or blatant failures. This chapter provides a general framework for the detailed discussions of the KFSH&RC data warehouse implementation discussed in the following chapters.

Nirvana

To early sailors, mermaids were half-fish, half-women that could lure boats onto the rocks and certain peril. In reality, of course, mermaids were nothing more than manatees. Similarly, to the uninitiated IT professional, the term *data warehouse* often conjures up a vision of a large hard-drive array that holds files from various hospital applications. In reality, a data warehouse is much more. For example, it is comprised of myriad technologies and processes that address three issues in the healthcare enterprise: transparency, standards, and adaptability/performance. To appreciate the significance of these issues, consider the ideal case.

In the future, healthcare CIOs will be able to select from a variety of shrink-wrapped hospital information systems (HISs). Unlike contemporary, disparate HISs, those in the future will run on the same hardware of choice; use a single clinical, administrative, and financial vocabulary; and support the real-time, graphical, and textual reporting of a virtually unlimited number of performance or quality indicators—all without degrading transaction performance. Systems will automatically compile statistics to demonstrate compliance with meaningful-use guidelines and seamlessly integrate with regional and national genomics and proteomics research centers.

Similarly, consider what will happen in this future scenario when the CIO receives a request for a new application—say, a new patient-tracking application based on implantable radio frequency identification (RFID) chips. After some financial and administrative maneuvering, the CIO will give the go-ahead to the development team. A programmer, perhaps offshore, will remotely drag and drop icons from a preconfigured object library. The resulting application will share the same database, vocabulary validation routines, and reporting capabilities as every other application in the system. Creating new printed reports or graphical dashboards will be a cinch, thanks to open, transparent, and documented architecture and database fields.

Despite the introduction of data warehouse appliances, open-source data warehouses, and innovations such as in-memory data warehouses, CIOs and domestic programmers in the healthcare industry need not worry about being replaced by shrink-wrapped HIS application suites any time soon. As of Q3 2016, only 4.5% of hospitals in the United States achieved Stage 7 on the HIMSS Analytics EMR Adoption Model, 30.5 % reached Stage 6, and just over one-third achieved Stage 5.[3] Furthermore, sizeable hospital information systems in the United States typically maintain applications on multiple operating systems and hardware platforms, and IT staff frequently deal with "rogue" departmental applications that require special care and handling, such as support for a legacy operating system or closed database engine.

There are also systemic issues in the healthcare environment beyond the control of the CIO. For example, there is no universal medical identifier in the United States. A seemingly obvious candidate, the Social Security number, is inherently flawed. One of several limitations is that many older women were never issued Social Security cards because it was assumed at one time that women would never work. There are also basic business issues, such as equating proprietary with profit. Consider that no one has been able to convince the likes of Cerner, Epic, GE, or Siemens AG to open up their proprietary databases and system architectures to facilitate integration with third-party applications.

Most hospital information systems are a confederation of variably dependent applications. As such, checking the heartbeat of an information system typically involves multiple queries against multiple systems, often involving mismatched patient identifiers. And then there is the issue of time. Days—and sometimes weeks—are often required to generate and validate complex reports that involve clinical, administrative, and financial data. Such poor performance would not be tolerated on Wall Street or in a typical Fortune 500 company, where time is money. However, it's the norm in healthcare.

As CIO, you have several options to address the transparency, standards, and performance limitations of a typical HIS. Interapplication interfaces, such as HL7, partially address these three issues, but they are generally limited in flexibility and in the number of data elements that can be shared among applications. Products such as SAS (www.sas.com) and SAP Crystal Solutions (www.sap.com) may be more viable solutions for a smaller healthcare organization with limited information systems resources. Another approach is to build a system from scratch, but this takes years and deep pockets, and results in—at best—another HIS standard to add to the endless list of standards.

Properly implemented, a data warehouse can provide the transparency, adherence to standards, and adaptability in our perfect system of the future. To the extent that a healthcare IT shop has access to and documentation on the database underlying the data warehouse, there is transparency. Constructing new reports is a matter of locating the relevant parameters in the documented database and manipulating it with appropriate reporting tools. Standards, including vocabulary, definitions, data structures, and operating systems, are integral to the design of a data warehouse. Similarly, performance, in terms of minimizing both query response time and the effect on transactional applications, is a feature of the properly implemented data warehouse.

A data warehouse, like the other options available to CIOs, has both benefits and liabilities. As typically implemented, a significant limitation of the data warehouse is that it serves as the basis for reporting and decision support, but not for transactional applications. In other words, the transparency is primarily useful for lightweight decision support applications that feed on the data warehouse. Transaction-based applications that must both read and write to a database are not supported.

Furthermore, standards must be selected carefully during the design of the data warehouse to avoid a confusing mix of standards that hinders system maintenance and prevents direct comparison of the healthcare enterprise performance with national and international benchmarks. In addition, the performance of the typical data warehouse can only approximate real time, in that data are at best updated every quarter-hour. More commonly, however, it is updated every night to minimize the negative impact on the source data applications. And, to add to the list of unmet challenges, most healthcare enterprises have yet to even consider the implications of interfacing with genomic and proteomic data from regional, national, and international centers. Clearly, there will be no shortage of work for healthcare CIOs for the foreseeable future.

Evolutionary Pressures

The economic and legislative pressures on the healthcare enterprise to provide quality healthcare at lower cost and with fewer resources have intensified. As costs shift from third-party payers to patients, the business of healthcare has begun to look like business in any other industry. Hospitals that provide superior outcomes, contain costs, and maintain profitability will

thrive at the expense of less fit institutions. CEOs in healthcare organizations are becoming increasingly aware of quality and performance management initiatives that have had a positive effect on the bottom line of businesses in other industries.

Some of the pressure on the modern healthcare enterprise is from performance-promoting organizations such as the Joint Commission, Centers for Medicare and Medicaid Services (CMS), the Agency for Healthcare Research and Quality (AHRQ), and the International Organization for Standardization (ISO). Most of these organizations promote the use of key performance indicators (KPIs) to help management more effectively direct the use of their organization's resources, maximize patient safety, promote clinical best practices, and increase patient satisfaction.

The Joint Commission's ORYX initiative includes performance indicators as part of its accreditation process. The organization defines performance in terms of outcome parameters, including efficiency, appropriateness, availability, timeliness, effectiveness, continuity, safety, and respect for caring. Its international equivalent, Joint Commission International (JCI), promotes quality standards that reflect practices outside of the United States.

CMS offers certification to healthcare organizations with quality initiatives that are planned, systematic, comprehensive, and ongoing. Specific, predetermined indicators and benchmarks form the basis of CMS performance indicators. AHRQ indicators cover access, utilization, cost, effectiveness, safety, timeliness, and patient-centeredness. ISO offers a process that a performance management system can follow for implementation.

Technology

Fiscal responsibility and the pressure of continuous quality improvement for healthcare IT favors a move from a disorganized system of different software packages running on different, incompatible hardware and abiding by various protocols to a seamless, organized system—and this is where the data warehouse comes in. Although it is often mistaken for an overgrown reporting system, the ideal data warehouse is a central, homogenous repository of a carefully selected subset of data from disparate, often loosely integrated applications in an organization. By virtue of this organization, the repository supports rapid data mining, report generation, and decision support.

Implementing a data warehouse is a technical challenge on several fronts, from data capture and transfer to controlling data access and handling the disposal of data. Consider that data from computers, RFID readers, bar code readers and bedside monitors must be acquired and made accessible in a way that is timely, accurate, secure, and HIPAA compliant. Furthermore, raw indicator values must be processed, filtered, and formatted before decision makers can use them as key quality indicators.

To appreciate the technological considerations and challenges inherent in implementing a data warehouse, consider the smaller and simpler clinical data repositories and data marts. A clinical data repository is a structured, systematically collected storehouse of patient-specific clinical data. These data are usually mirrored from a single clinical application but may be supplemented with data from other clinical systems. By maintaining a separate database, configured specifically to support decision analysis, the application database engine is spared computational loading, and the response time to a particular query should be improved.

Furthermore, because virtually all patient information in the host application is mirrored in the clinical data repository, complex, customized queries are possible without degrading the performance of the source applications. In addition, because the data tend to originate from one source, with little to no data manipulation, near-real-time retrieval of clinical data is possible.

Stepping up one level of complexity, a data mart contains data extracted from clinical and nonclinical applications, including summary data. In operation, a select subset of data from multiple transactional applications are checked for errors, summarized, and imported into a central database. Data marts tend to be used at the department level and are often isolated from the larger healthcare enterprise.

A data warehouse is an enterprise-wide central repository of information that reflects activity within most applications running in the enterprise. As with a data mart, a data warehouse combines data from a variety of application databases into a central database. This requires cleaning, encoding, and translating data so that analysis can be performed. Data redundancy may be intentionally built in to the data warehouse to maximize the efficiency of the underlying database engine—for example, by minimizing the number of relational tables to be joined in a report query. There are also the usual database issues to consider, such as security, data integrity, synchronization, failure recovery, and general data management.

People

Technology is necessary but insufficient for the continued evolution of the healthcare enterprise. As with prepping a seagoing vessel for a long journey, a well-provisioned and trained staff goes a long way to mitigate the risk of failure. This reliance on trained staff is recognized by the federal incentive funding under the HITECH component of the American Recovery and Reinvestment Act of 2009; meaningful-use criteria are focused on people-oriented organizational change, not technology. Creating a hospital information system that is actually used requires all stakeholders to understand the mission of the enterprise, share the vision of the administration, and have the motivation to overcome the challenges that must be addressed. However, even the best-intentioned CIO or hospital administrator is powerless to make the appropriate change without timely, accurate, and relevant information.

As many medical IT professionals discovered long ago, any technological enablers must be embraced by the user communities for the technologies to have a positive impact. Simply providing decision makers with a torrent of data through sophisticated, hi-tech graphical displays is worthless without an underlying strategy.

One such strategy is performance management. The basis of performance management is the effective use of resources, as measured by quantifying processes and outcomes using KPIs that gauge the performance of an organization in particular areas. Performance management initiatives that have been applied in healthcare and other industries include aspects of statistical process control, total quality management, customer relationship management, activity-based costing, ISO 9000/ISO 9001:2015, and knowledge management. Because performance management is a tough sell in clinically based organizations, initiatives are often better defined in terms of quality.

Knowledge management—a deliberate, systematic business optimization strategy that involves the selection, distillation, storage, organization, packaging, and communication of information—is particularly relevant to the success of a data warehouse–enabled performance management project. This strategy treats intellectual capital, including process, structures, information systems, financial relations, and intellectual properties, as a major organizational asset that can be tracked, measured, and analyzed with performance or quality indicators. (See Figure 1.3 for a map of typical knowledge management operations.) Knowledge management is practiced to some degree

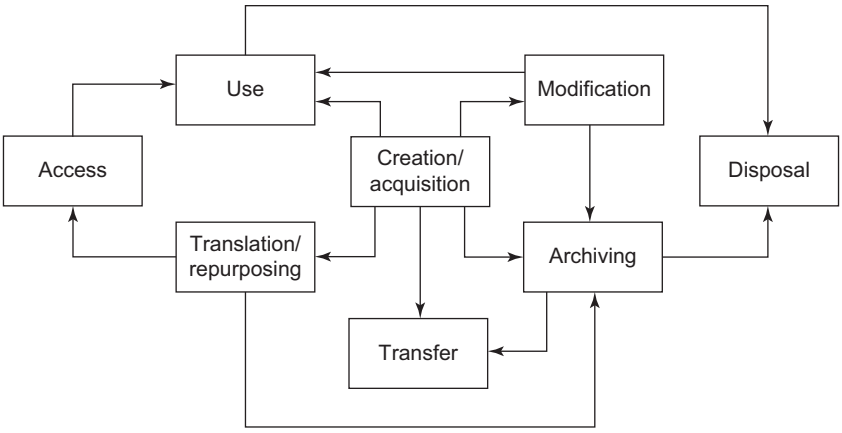

Figure 1.3 Typical knowledge management operations.

in every successful knowledge-intensive organization, including the data-driven healthcare enterprise.

Monsters

Yes, there be monsters here. However, just as the sea captains of old went forth in spite of the risks, healthcare CIOs are moving forward with data warehouse implementations. Those that succeed are prepared. They both know exactly where they're going and have a good idea of how to get there. For example, the successful CIO knows that the ideal data warehouse automatically downloads data from application databases, cleans and transforms data as necessary, and then combines them into a central database. A properly implemented data warehouse also takes care of timing issues and populates a central database in such a way as to support the most likely queries to be asked. Ideally, most data warehouse performance management efforts are begun through a lengthy requirements process; all the key users select the fields that are used to populate their most used queries. These fields are voted on before the data warehouse is built.

The successful CIO knows that most of the technical implementation challenges are related to the independently designed application database systems that rely on different data representations, unique vocabularies, and different update timings. For example, one application might represent date as "day/month/year," whereas another application uses "month/day/year." In order for the data warehouse to provide valid date information, data

from one or more application databases must be translated into the representation used in the data warehouse. Only then can the data be sorted, massaged, translated, and reformatted to support data mining, discovering patterns in the data, compiling outcome statistics, or performing ad hoc queries.

Successful CIOs are also aware that variation in application update timings creates data warehouse timing and synchronization challenges. Ideally, all information entering the data warehouse represents an instant in time when all transactions are frozen and data edits and modifications are halted. In reality, even if the data are downloaded from each transactional database at the same instant, they may be out of sync because of how the applications are written. For example, one application may write data out to disc every hour, whereas a second application writes data to disc immediately after each transaction.

Variations in application vocabularies present unique challenges as well. A central issue in data warehouse design is that there are several vocabulary standards available for use in the central database and query engines. SNOMED, DICOM, ICD-10, and UMLS all have issues related to completeness and applicability to particular clinical and nonclinical domains.

In assessing the many challenges associated with implementing a data warehouse, it is tempting to focus on the technology. After all, technology is logical, controllable, and eventually works, given sufficient time and effort. However, as you'll note in the accounts contained in the following chapters, the greatest challenges are related to people, not technology. The success of any data warehouse implementation will be limited to the degree that your people and the processes are in place to work with the system. Because there will always be doubts in the minds of the men and women who do the heavy lifting, leadership, whether taking your team across uncharted waters or through a data warehouse implementation, is the greatest determinant of success.

References

1. Gartner 2005. Gartner Says More than 50 Percent of Data Warehouse Projects Will Have Limited Acceptance or Will Be Failures through 2007. Gartner Newsroom. February 24, 2005. Available at: http://www.gartner.com/newsroom/id/492112.