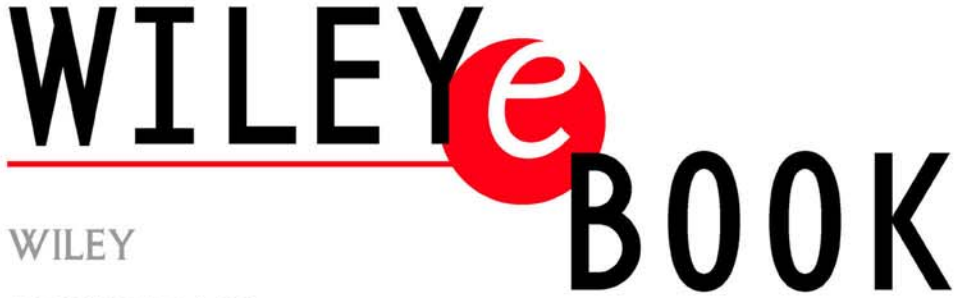
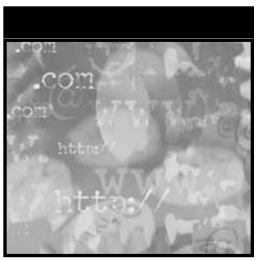
Business Culinary Architecture
Computer General Interest
Children Life Sciences Biography
Accounting Finance Mathematics
History Self-Improvement Health
Engineering Graphic Design
Applied Sciences Psychology
Interior Design Biology Chemistry



WILEY
JOSSEY-BASS
PFEIFFER
J.K.LASSER
CAPSTONE
WILEY-LISS
WILEY-VCH
WILEY-INTERSCIENCE



Common Warehouse Metamodel

John Poole
Dan Chang
Douglas Tolbert
David Mellor

Wiley Computer Publishing

John Wiley & Sons, Inc.

NEW YORK • CHICHESTER • WEINHEIM • BRISBANE • SINGAPORE • TORONTO

Publisher: Robert Ipsen
Editor: Robert Elliott
Assistant Editor: Emilie Herman
Managing Editor: John Atkins
Associate New Media Editor: Brian Snapp
Text Design & Composition: D&G Limited, LLC

Published by John Wiley & Sons, Inc., New York

This title is also available in print as ISBN 0-471-20052-2. Some content that appears in the print version of this book may not be available in this electronic edition.

For more information about Wiley products, visit our web site at www.Wiley.com

To my Mother, Josephine F. Poole, for her love and encouragement, and for teaching me the value of science at an early age, and to the memory of my father, John R. Poole, a true man among men.

—John

To my wife, Hedy, and my daughters, Lillian, Anne, and Vivian, who are the best in my life.

—Dan

To Margaret and Lindy.

—Doug

To my wife Michele, for her boundless love, support, and encouragement, and to the memory of my father, Richard G. Mellor.

—David

Advance Praise for *Common Warehouse Metamodel*

"Business intelligence applications have become the new data processing 'stove pipes' within the departments of an organization. In order for the analytics of these solutions to benefit the entire enterprise, standards such as CWM are essential. This book is an outstanding primer on CWM, not only giving the technical foundation behind this standard, but detailing the value of its application."

John Kopcke
Chief Technology Officer, Hyperion Solutions Corporation

"The CWM standard is one of the most important standards in the business intelligence and meta data repository market segments. Anyone in the business of building a meta data repository—whether for a software product or a corporate repository—needs to own a copy of this book."

David Marco
President, Enterprise Warehousing Solutions, Inc.

"For all but the smallest of organizations, high-quality business intelligence or decision support requires the integration of multiple information technologies—such as relational and multidimensional databases, data mining, and visualization—within a domain-specific solution-oriented framework. Unfortunately, the disparity of implicit modeling frameworks and meta data between different types of analytical technologies, not to mention between different vendors' products of the same type, has kept

the benefits of integration beyond all but a few of the most price-insensitive applications. The OMG's Common Warehouse Metamodel represents an important step in the direction of multi-technology integration for enterprise decision support.

This book provides a timely, well-organized, clearly written, and comprehensive overview of CWM. I highly recommend it for practitioners who need to build CWM-based solutions, data-warehousing and decision support managers who need to understand the changing landscape of technology options, and for anyone interested in standards initiatives for data warehousing and decision support."

Erik Thomsen
Chief Scientist, DSS Lab

"It is easy to dismiss the Common Warehouse Metamodel (CWM) specification as 'just another three-letter acronym.' The sometimes-arcane world of data modeling and information architecture can seem impenetrable at times. But the emergence of this single worldwide standard for data warehousing is critical to solving the problem every CIO has: leveraging existing information assets to deliver new goods and services. Poole, Chang, Tolbert, and Mellor provide the first comprehensive introduction and guidebook to this all-important specification in a clear, concise, and complete book. It belongs on the bookshelf of every information architect, data-modeling professional, and database administrator."

Dr. Richard Mark Soley
Chairman and CEO, Object Management Group, Inc.

"Now that information is regarded by leading organizations as the fourth pillar of enterprise assets (in addition to financial, material, and human capital), formal meta data management has emerged as a key enabler of improved information asset valuation. Managing and leveraging information simply cannot be accomplished effectively without meta data. Until this book, only abundant, redundant, and incoherent meta data theory has been available. This book's treatment incorporates the gritty specifics craved by practitioners and product developers. Based on the leading meta data standard, this text diametrically details the approach to implementing practicable meta data solutions as part of custom information management applications or commercial offerings."

Doug Laney
Vice President, Application Delivery Strategies (ADS), META Group

2001 OMG Press Advisory Board

OMG Press Books in Print

For complete information about current and upcoming titles, go to http://www.Wiley.com/compbooks/.

- ***Building Business Objects*** by Peter Eeles and Oliver Sims, ISBN: 0471-191760.
- ***Business Component Factory: A Comprehensive Overview of Component-Based Development for the Enterprise*** by Peter Herzum and Oliver Sims, ISBN: 0471-327603.
- ***Business Modeling with UML: Business Patterns at Work*** by Hans-Erik Eriksson and Magnus Penker, ISBN: 0471-295515.
- ***CORBA 3 Fundamentals and Programming, 2nd Edition*** by Jon Siegel, ISBN: 0471-295183.
- ***CORBA Design Patterns*** by Thomas J. Mowbray and Raphael C. Malveau, ISBN: 0471-158828.
- ***Enterprise Application Integration with CORBA: Component and Web-Based Solutions*** by Ron Zahavi, ISBN: 0471-32704.
- ***Enterprise Java with UML*** by CT Arrington, ISBN: 0471-386804
- ***Enterprise Security with EJB and CORBA*** by Bret Hartman, Donald J. Flinn and Konstantin Beznosov, ISBN: 0471-150762.
- ***The Essential CORBA: Systems Integration Using Distributed Objects*** by Thomas J. Mowbray and Ron Zahavi, ISBN: 0471-106119.
- ***Instant CORBA*** by Robert Orfali, Dan Harkey, and Jeri Edwards, ISBN: 0471-183334.
- ***Integrating CORBA and COM Applications*** by Michael Rosen and David Curtis, ISBN: 0471-198277.
- ***Java Programming with CORBA, Third Edition*** by Gerald Brose, Andreas Vogel, and Keith Duddy, ISBN: 0471-247650.
- ***The Object Technology Casebook: Lessons from Award-Winning Business Applications*** by Paul Harmon and William Morrisey, ISBN: 0471-147176.
- ***The Object Technology Revolution*** by Michael Guttman and Jason Matthews, ISBN: 0471-606790.
- ***Programming with Enterprise JavaBeans, JTS and OTS: Building Distributed Transactions with Java and C++*** by Andreas Vogel and Madhavan Rangarao, ISBN: 0471-319724.
- ***Programming with Java IDL*** by Geoffrey Lewis, Steven Barber, and Ellen Siegel, ISBN: 0471-247979.
- ***Quick CORBA 3*** by Jon Siegel, ISBN: 0471-389358.
- ***UML Toolkit*** by Hans-Erik Eriksson and Magnus Penker, ISBN: 0471-191612.

About the OMG

The Object Management Group (OMG) was chartered to create and foster a component-based software marketplace through the standardization and promotion of object-oriented software. To achieve this goal, the OMG specifies open standards for every aspect of distributed object computing from analysis and design, through infrastructure, to application objects and components.

The well-established CORBA (Common Object Request Broker Architecture) standardizes a platform- and programming-language-independent distributed object computing environment. It is based on OMG/ISO Interface Definition Language (OMG IDL) and the Internet Inter-ORB Protocol (IIOP). Now recognized as a mature technology, CORBA is represented on the marketplace by well over 70 ORBs (Object Request Brokers) plus hundreds of other products. Although most of these ORBs are tuned for general use, others are specialized for real-time or embedded applications, or built into transaction processing systems where they provide scalability, high throughput, and reliability. Of the thousands of live, mission-critical CORBA applications in use today around the world, over 300 are documented on the OMG's success-story Web pages at www.corba.org.

CORBA 3, the OMG's latest release, adds a Component Model, quality-of-service control, a messaging invocation model, and tightened integration with the Internet, Enterprise Java Beans, and the Java programming language.

Widely anticipated by the industry, CORBA 3 keeps this established architecture in the forefront of distributed computing, as will a new OMG specification integrating CORBA with XML. Well-known for its ability to integrate legacy systems into your network, along with the wide variety of heterogeneous hardware and software on the market today, CORBA enters the new millennium prepared to integrate the technologies on the horizon.

Augmenting this core infrastructure are the CORBAservices, which standardize naming and directory services, event handling, transaction processing, security, and other functions. Building on this firm foundation, OMG Domain Facilities standardize common objects throughout the supply and service chains in industries such as Telecommunications, Healthcare, Manufacturing, Transportation, Finance/Insurance, Electronic Commerce, Life Science, and Utilities.

The OMG standards extend beyond programming. OMG Specifications for analysis and design include the Unified Modeling Language (UML), the repository standard Meta-Object Facility (MOF), and XML-based Metadata Interchange (XMI). The UML is a result of fusing the concepts of the world's most prominent methodologists. Adopted as an OMG specification in 1997, it represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems and is a well-defined, widely accepted response to these business needs. The MOF is OMG's standard for metamodeling and meta data repositories. Fully integrated with UML, it uses the UML notation to describe repository metamodels. Extending this work, the XMI standard enables the exchange of objects defined using UML and the MOF. XMI can generate XML Data Type Definitions for any service specification that includes a normative, MOF-based metamodel.

In summary, the OMG provides the computing industry with an open, vendor-neutral, proven process for establishing and promoting standards. OMG makes all of its specifications available without charge from its Web site, www.omg.org. With over a decade of standard-making and consensus-building experience, OMG now counts about 800 companies as members. Delegates from these companies convene at week-long meetings held five times each year at varying sites around the world, to advance OMG technologies. The OMG welcomes guests to their meetings; for an invitation, send your email request to info@omg.org.

Membership in the OMG is open to end users, government organizations, academia, and technology vendors. For more information on the OMG, contact OMG headquarters by phone at 1-508-820-4300, by fax at 1-508-820-4303, by email at info@omg.org, or on the Web at www.omg.org.

Contents

Foreword

The world of meta data began with simple listings that came from early compilers. From that humble origin came directories, then data dictionaries, then repositories and sophisticated end-user tools. The Internet came along and opened up the doors of computing into venues never before imagined, and meta data grew once again.

The corporate information factory (CIF) gave these different forms of computing a cohesive framework. From the data warehouse–centric CIF came the ODS, data marts, DSS, applications, exploration warehouses, data mining warehouses, and alternate storage. They required coordination across the framework. It naturally fell to meta data to provide the means for communication across the different components of the CIF.

The need for meta data began with simple documentation. System designers and developers, and later data modelers, found they needed meta data. Next systems integration consultants discovered that meta data was the glue that held things together. Finally the architect of the CIF came to the conclusion that

without meta data, the CIF was just a bunch of architectural components operating independently and in chaos. Without meta data, one architectural component was singing rock and roll, another architectural component was singing rhythm and blues, and yet another component was singing country and western. The resulting cacophony was hardly pleasing to the ear. To get the components singing together, meta data across the CIF was essential.

Today the world of meta data has extended well beyond the simple technical origins of data dictionaries into the world of distributed processing, business meta data, and the Internet. Yet there is still surprising resistance to the standardization of meta data: From vendors who are determined not to let their meta data fall into the hands of other vendors, resulting in an industry full of proprietary meta data without standardization; From venture capitalists who prefer dot-com opportunities to financing new meta data projects; and from academics who are so far removed from the real issues of the day that any practical approach to meta data seems to elude them.

The time has come for consumers to step forward and specify a standard for meta data semantics and exchange. This book is an important step toward that goal. It will be a lodestone for future products and collaborative efforts across the corporate information factory.

W.H. Inmon
www.billinmon.com

Acknowledgments

The Common Warehouse Metamodel specification would never have been possible without the efforts and contributions of many individuals. The authors wish to acknowledge and thank their many colleagues who participated in, contributed materially to, reviewed, and strongly supported, the CWM specification. This includes (in alphabetical order):

Steve Allman—Dimension EDI

Don Baisley—Unisys Corporation

Christian Bremeau—Meta Integration Technology, Inc.

Gorden Callen—Oracle Corporation

Cory B. Casanave—Data Access Technologies

J. J. Daudenarde—IBM Corporation

David S. Frankel—Genesis Development Corporation

Jack J. Greenfield—InLine Software

Tina Groves—Cognos, Inc.

Hans-Peter Hoidn—UBS AG

Mark Hornick—Oracle Corporation

Sridhar Iyengar—Unisys Corporation

David Last—Oracle Corporation

Debra LaVergne—IBM Corporation

Christoph Lingenfelder—IBM Corporation

Phil Longden—Genesis Development Corporation

Chuck Mosher—Sun Microsystems, Inc.

Jishnu Mukerji—Hewlett-Package Company

Chris Nelson—Dimension EDI

Robin Noble-Thomas—IBM Corporation

Jeffrey Peckham—UBS AG

Pete Rivett—Adaptive Ltd.

Karsten Reimer—Sun Microsystems, Inc.

Vilhelm Rosenqvist—NCR Corporation

Charles E. Simon—Aonix

David C. Smith—Deere & Company

Anders Tornqvist—Dimension EDI

Barbara Walters—SAS Institute, Inc.

Yuichi Yagawa—Hitachi, Ltd.

David Zhang—Hyperion Solutions Corporation

The authors would also like to express their gratitude to the following individuals who provided considerable support and backing to the CWM specification effort:

Ravi Dirckze, Susan Donahue, Giuseppe Facchetti, Robert Kemper, Suresh Kumar, James Jonas, Joanne Lamb, Don Lind, Tony Maresco, Bruce McLean, Karel Pagrach, William Perlman, Jeff Pinard, Curtis Sojka, Chris de Vaney, Robert Vavra, Adriaan Veldhuisen.

And, of course, tremendous thanks go to the fine editorial staff at John Wiley & Sons, who recognized the importance of this book and made its publication possible:

John Atkins, Julie Frigo, Robert M. Elliott, Emilie Herman, Theresa Hudson, and Kelly D. Dobbs.

Introduction to the Common Warehouse Metamodel

Meta data: data describing data, or information about data; generally comprised of descriptions of information structures. Meta data is critical to achieving integration between dissimilar software systems and products from multiple vendors. If heterogeneous software products, applications, reporting tools, and databases are to interoperate effectively, they must have a common understanding of each other's information structures, which requires a common language in which to describe and communicate meta data. They must also have standard interfaces and protocols that support meta data interchange.

This fact is perhaps most relevant in the data warehousing and business analysis domains, in which the various components comprising the *information supply chain* (for example, operational data stores, relational databases, OLAP servers, analysis, visualization, and reporting tools) possess rich meta data structures. These components rely heavily on this meta data as the foundation of their operations and services. However, it is nearly impossible for most commercial software products and systems to readily share meta data. Most products coming from different vendors have dissimilar meta data models (or *metamodels*) and proprietary interfaces exposing their meta data. These differences result in considerable cost and diminished return-on-investment to both vendors and customer organizations attempting to integrate products, tools, and applications.

Nearly all integration efforts require the construction of custom *bridges* between dissimilar metamodels. This construction is time consuming, costly, and results in the creation of many software modules that perform essentially the same function but cannot readily be reused in other integration efforts.

As data warehousing becomes increasingly Web-centric (for example, using data warehouses to store and analyze clickstreams for online retailing and other e-commerce applications), an even greater requirement exists for robust meta data-level integration, especially as data warehouse components are deployed in totally heterogeneous, collaborative, and distributed application environments. The current climate of merger and acquisition, as well as the general desire to avoid single vendor lock-in and to combine best-of-breed products in the building of an information system, also places considerable pressure on vendors to provide readily interoperable tools, products, and applications.

The CWM Solution

The Common Warehouse Metamodel (CWM) is a recently adopted standard of the Object Management Group (OMG) for meta data interchange in the data warehousing and business analysis environments. CWM provides the long sought-after common language for describing meta data (based on a generic, but semantically replete, common data warehousing and business analysis domain metamodel) and an XML-based meta data interchange facility. CWM is rapidly gaining momentum within the data warehousing and business analysis communities and is being incorporated into various vendors' next generation of data warehousing products and tools.

Now that a single, industry-wide standard for meta data interchange exists, vendors finally have the common metamodel and interchange mechanism they need to build truly interoperable databases, tools, and applications. Customers will benefit by being able to select from best-of-breed product offerings while remaining confident that their investments will not be diluted by the inability of tools to interoperate.

From a technical standpoint, CWM extends the OMG's established metamodeling architecture to include data warehousing and business analysis domain concepts. CWM supports a model-driven approach to meta data interchange, in which formal models representing shared meta data are constructed according to the specifications of the CWM metamodel (essentially an object technology approach to achieving data warehouse integration). These models are stored and interchanged in the form of XML documents. Meta data can be defined independently of any product-specific considerations or formats. It can be stored externally to products as an information commodity

within its own right and is readily used by products as generic definitions of information structures.

Data warehousing and business analysis tools that agree on the fundamental domain concepts and relationships defined by CWM can understand a wide range of models representing particular meta data instances. Tools, products, and applications can integrate at the meta data level, because they have a common language in which to externalize their meta data and do not require knowledge of each other's proprietary information structures and interfaces.

Mission of This Book

The mission of this book is to provide a single, coherent, and comprehensive overview of the OMG's Common Warehouse Metamodel, which is easy to read. Prior to this book's publication, no such single source of information on CWM was readily available. Nearly all of the available information on CWM was scattered over the Internet, largely comprised of the CWM specifications themselves, numerous presentation slides, position papers, press kits, analyst reports, various summaries, and overviews. Readers needed to sift through a prohibitively large and rather fragmentary collection of information to gain even a rudimentary understanding of CWM and the value it provides to both vendors and their customers.

The intent of the book is basically threefold:

- Fill the aforementioned information gap with a concise and highly readable primer on CWM, written by several of the core developers of the CWM standard.
- Build interest, awareness, and understanding of CWM and its value in the data warehousing and business analysis marketplaces. The authors want to ensure that broad adoption of CWM is not inhibited by a dearth of reader-friendly and sensibly organized information on CWM.
- Serve as an introduction and overview for the forthcoming *Common Warehouse Metamodel Developer Guide*, published by John Wiley & Sons, which will serve practitioners who are attempting to implement CWM within their own products or deploy CWM-enabled tools within their own organizations and data warehouses.

Perhaps the most ambitious goal of this book is to impart on the reader a CWM conceptual understanding or worldview comparable to that possessed by those individuals who participated in the actual development and implementation of CWM. This level of understanding is critical for those who want

to deploy or use CWM in achieving meta data-based interoperability and integration in data warehousing and business analysis projects.

How This Book Is Organized

The organization of this book introduces the reader to the Common Warehouse Metamodel in a reasonably straightforward and logical manner. Each chapter builds upon preceding chapters and chapters should generally be read in order, although readers generally familiar with the problems of meta data integration and wanting to get immediately into the details of CWM may safely skip the first two chapters and start with Chapter 3, "Foundation Technologies."

Chapter 2, "The Value Proposition for CWM," introduces the overall problems associated with attempts at integrating the products and tools comprising the information supply chain in data warehousing environments. This chapter details the value proposition for CWM—that is, why a CWM-based solution to data warehouse integration enhances the overall return-on-investment (ROI) of the data warehouse—and provides a detailed overview of the model-driven approach to integration espoused by CWM.

Chapter 3, "Foundation Technologies," provides an overview of the various foundational technologies of the OMG's metamodeling architecture that are leveraged by CWM. These technologies include the OMG's Meta Object Facility (MOF), Unified Modeling Language (UML), XML Meta Data Interchange (XMI), and Interface Definition Language (IDL). The Java programming language also is discussed as the implementation language of choice for most CWM implementation efforts.

Chapter 4, "An Architectural Overview of CWM," continues the discussion by examining how CWM extends the OMG metamodeling architecture into the data warehousing and business analysis domains. CWM is notably the first OMG standard to extend this architecture into a specific problem domain. A detailed survey of CWM's own layered architecture is presented, including CWM core components and extension packages.

Concrete examples of how CWM is used to model various real-world data warehousing situations are presented in Chapter 5, "Using CWM." Descriptions of how CWM relates to several of the more traditional data warehousing models are provided, along with general guidelines for constructing CWM models.

The planning, design, implementation, and deployment of CWM within the context of data warehousing architectures and cohesive meta data management strategies are described throughout Chapter 6, "Developing Meta Data Solutions Using CWM," and Chapter 7, "Implementing CWM." Collectively,

these chapters unite much of the higher level, business value proposition discussions presented in Chapter 2 with the more detailed, technical information presented in Chapters 3 through 5.

Finally, Chapter 8, "Conclusions," discusses how CWM relates to established and forthcoming standards for system integration and interoperability, such as CORBA, XML, and Web-oriented data-interchange protocols. It provides speculation on the technical evolution of the CWM standard and recommendations for further reading and research.

Who Should Read This Book

This book is intended for a moderately technical audience. The majority of our readers will be software developers, system integrators, software and data warehouse architects, database administrators, technical managers, product managers, marketing managers, planners, and strategists working primarily in the data warehousing and business analysis domains. The book will be of benefit to anyone who needs to rapidly build a coherent CWM knowledge base.

Technology-aware, high-level administrators, executives, and other decision-makers responsible for the overall management and integration of their organization's data warehouses, data warehousing, business analysis product offerings, and corporate-wide meta data strategies will find this book highly useful. Academicians and engineers from other disciplines who are interested in learning more about standards-based data warehousing and integration should also find this book very useful.

What's on the Web Site

This book is accompanied by the Web site www.wiley.com/compbooks/poole. This free site provides updated information on the ongoing evolution of the CWM standard as it affects the content and goals of the book. Also provided are links to vendors offering CWM-enabled data warehousing and business analysis tools, as well as links to other active or developing sources of information on CWM.

A Brief History of CWM

A considerable amount of history in the field of meta data standardization and integration preceded the development and adoption of CWM, and the following discussion attempts to summarize much of that history.

The OMG's adoption of the UML as a standard in 1997 is the first significant event leading to the genesis of the CWM. UML is a formal language for modeling discrete systems. Although most people familiar with UML tend to view it as a visual modeling language for designing object-oriented software systems, UML is not necessarily tied to this particular usage. Because UML is a formally defined language, artifacts modeled in UML (for example, descriptions of data warehousing concepts, in the case of CWM) are easily translated to other notations (such as XML), which are not necessarily visual in nature. These other notations may provide other benefits, however, including the easy communication of model contents between software processes.

The second major antecedent event to the creation of CWM was the adoption of the MOF as an OMG standard, also in 1997. MOF defines an abstract syntax for describing models of meta data (that is, metamodels). MOF is based on a subset of the UML and can describe metamodels such as the UML and CWM metamodels and even the MOF itself. In practical terms, MOF defines how meta data repositories might be structured and what capabilities and interfaces repositories need to provide to ensure that all supported metamodels have a common linguistic framework.

XMI was the third major cornerstone technology paving the way for CWM. The XMI specification defines how MOF-based meta data is represented in XML. Meta data is stored in an XML document, and XML tags provide definitions of the meta data elements. XMI is a significant foundational technology for CWM, because it provides a low-cost and Web-enabled interchange mechanism for meta data. XMI was adopted by the OMG as a standard early in 1999.

While the foundations for CWM were being developed within the OMG, several other significant meta data standardization efforts and trends were taking place in the industry. As early as 1993, the Electronics Information Group published the CASE Data Interchange Format (CDIF), a standard for interchanging meta data generated from Computer Aided Software Engineering (CASE) tools. CDIF gained some degree of industry acceptance but had perhaps arrived on the scene at the wrong time (after CASE tools had begun to fall out of the industry's favor but prior to the meta data efforts that started in the mid-to-late 1990s).

In October 1995, the Meta Data Coalition (MDC) was formed by a number of IT industry leaders. In April 1996, the MDC released the Meta Data Interchange Specification (MDIS) Version 1.0. MDIS was a product-neutral mechanism for the interchange of meta data between MDIS-enabled tools and consisted of a metamodel, an access framework specification, and a tag-oriented language for specifying meta data instances. In retrospect, MDIS was considerably narrower in scope than CWM, focusing primarily on the interchange of meta data-defining database schemas. Its specification and interchange languages

were also proprietary to MDIS, in contrast to the use of UML and XML in the CWM (UML and XML had not been widely received by the industry when MDIS was drafted). MDIS was a note-worthy attempt, however, and represents one of the first *linguistic* (that is, metamodel and tag language-oriented) approaches to the meta data interchange problem.

At about the same time MDIS was being developed by the MDC, Microsoft Corporation was developing the Open Information Model (OIM), along with several other contributors, and in October 1996, an early draft of OIM was made available. What was significant about OIM was that it leveraged UML as its specification language and was based heavily on repository technologies developed largely by Microsoft and other contributors.

About two years later, several significant events occurred that brought about the recent competition between emerging meta data standards in the IT industry. In September 1998, the OMG published its RFP for a *Common Warehouse Meta Data Interchange* specification that was to build upon existing OMG technology standards for meta data and repository architectures (that is, UML, MOF, and XMI). In early 1999, several OMG member organizations (primarily IBM, Unisys, and Oracle) decided to collaborate on the joint submission of a proposal in response to this RFP. An intercompany team comprised of software architects from IBM, Unisys, NCR, Hyperion, Oracle, UBS AG, Genesis, and Dimension EDI set about developing this metamodel-based solution, under the leadership of Dr. Daniel T. Chang of IBM's Silicon Valley Laboratories. This effort produced what subsequently became known as the OMG's *Common Warehouse Metamodel* (CWM).

Meanwhile, in December 1998, Microsoft joined the MDC and submitted OIM to the MDC, ensuring that OIM would be developed further as an open standard. Representatives from the various MDC member companies subsequently released OIM Version 1.0 in July 1999. The subsequent submission of the initial CWM specification to the OMG occurred in September 1999, and this was immediately followed by an interoperability demonstration of CWM technology by several of the cosubmitting organizations that following November. It became readily apparent that the industry was now faced with the prospect of two competing and closely related specifications for meta data integration. This situation was met with considerable concern. How could true meta data integration be achieved if different products embraced different standards? Were the two standards reasonably similar to one another or highly divergent and fundamentally incompatible? Could the two standards somehow be merged, or otherwise bridged?

The OIM, like CWM, was based heavily on UML and relied on an XML-encoding of its meta data to provide a meta data interchange format. The OIM metamodel represented enterprise-wide IT architectures, including elements representing the system analysis and design process, component-based

deployment, database, data warehousing, and knowledge management. In many ways, the scope of OIM was broader than that of CWM, focusing on the overall IT environment; whereas CWM focused specifically on data warehousing and business intelligence. Where the two standards converged (that is, data warehouse architecture and business analysis), CWM provided complete coverage of OIM semantics. This was due, in part, to the fact that OIM had served as one of the major design references in the development of CWM. The other reference standards included CDIF, MDIS Version 1.1, which was subsequently absorbed by OIM Version 1.0, and the OLAP Council's Multidimensional MD-API Version 2.0.

Therefore, where CWM and OIM intersected, CWM supported all of the capabilities of OIM and possessed an open standards-based repository infrastructure and meta data interchange mechanism. Those areas where OIM and CWM diverge generally reside outside the scope of data warehousing and business analysis. However, the most promising area for meta data-based ROI currently resides more in the area of data warehousing and business analysis tool integration and less in the broader areas of enterprise-wide environments. Therefore, although the CWM solution was narrower in overall scope than OIM, it had greater focus on a solution space that promised a much more immediate ROI and a greater probability of near-term success. It offered more bang for the buck (more on this in Chapter 2).

As a first step toward ensuring some level of convergence between the two standards, the OMG and MDC established reciprocal membership and representation between both organizations. An ongoing dialog took place, along with sharing of information and work products. It was even proposed that the two standards could possibly be bridged at the interchange level using standard XML transformations, and a proof-of-concept demonstration of this approach was performed.

However, in June 2000, the OMG Board of Directors approved the second edition of the CWM as an OMG *Adopted Technology Specification*. That September, given the support for CWM building within the industry, the MDC membership voted to discontinue its efforts on OIM in favor of joining ranks with the OMG and focusing on the continued development of the CWM standard. From this point on, the industry finally had a single, open standard for meta data integration with broad vendor support.

Much of the subsequent effort on CWM consisted of initial product-level implementations of the CWM standard. This culminated in the presentation of a multivendor *CWM Enablement Demonstration* at the OMG Technical Conference in December 2000. The experiences gained from initial product-level implementations were used to further tune the specification, leading to the submission of CWM Version 1.0 to the OMG. CWM Version 1.0 was approved by the OMG membership in March 2000, and this was immediately followed

by another multivendor interoperability demo at the Meta Data Conference/DAMA Symposium, also in March. The OMG Board of Directors formally adopted CWM Version 1.0 as an OMG *Available Technology Specification* the following month, during the April OMG Technical Conference in Paris.

As part of the CWM adoption process, a CWM Revision Task Force has been formed to take ownership of CWM and further the evolution of the CWM standard through the formal OMG revision process. The membership of the current CWM RTF includes a number of previous supporters and developers of OIM.