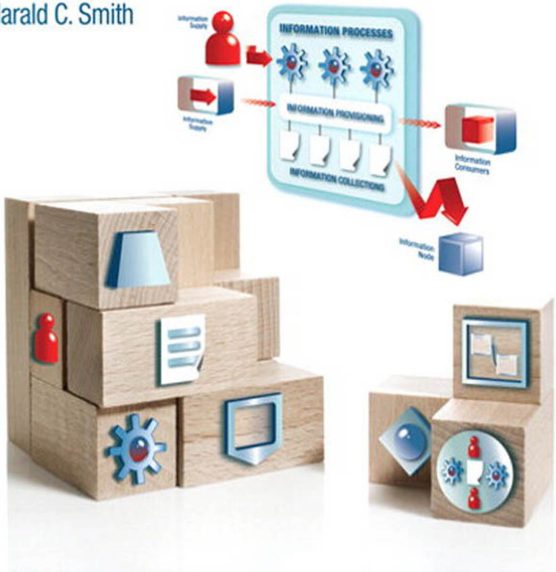
Patterns of Information Management

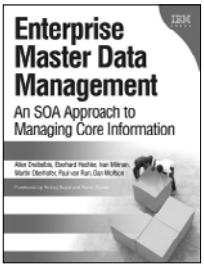
Mandy Chessell
Harald C. Smith



FREE SAMPLE CHAPTER

SHARE WITH OTHERS

Related Books of Interest



Enterprise Master Data Management

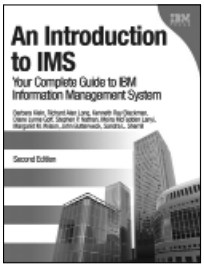
An SOA Approach to Managing Core Information

By Dreibelbis, Hechler, Milman, Oberhofer, van Run, Wolfson

ISBN-13: 978-0-13-236625-0

The Only Complete Technical Primer for MDM Planners, Architects, and Implementers

An authoritative, vendor-independent MDM technical reference for practitioners: architects, technical analysts, consultants, solution designers, and senior IT decision makers. Written by the IBM® data management innovators who are pioneering MDM, this book systematically introduces MDM's key concepts and technical themes, explains its business case, and illuminates how it interrelates with and enables SOA.



An Introduction to IMS

Your Complete Guide to IBM Information Management Systems, 2nd Edition

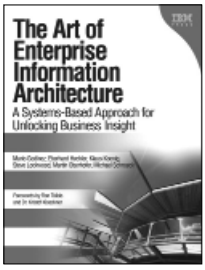
By Barbara Klein, et al.

ISBN-13: 978-0-13-288687-1

IBM's Definitive One-Stop Guide to IMS Versions 12, 11, and 10: for Every IMS DBA, Developer, and System Programmer

Over 90% of the top Fortune® 1000 companies rely on IBM's Information Management System (IMS) for their most critical IBM System z® data management needs: 50,000,000,000+ transactions run through IMS databases every day. What's more, IBM continues to upgrade IMS: Versions 12, 11, and 10 meet today's business challenges more flexibly and at a lower cost than ever before. In ***An Introduction to IMS, Second Edition***, leading IBM experts present the definitive technical introduction to these versions of IMS.

Related Books of Interest



The Art of Enterprise Information Architecture

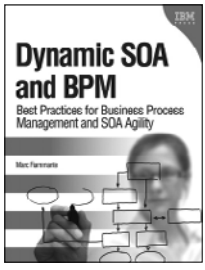
By Godinez, Hechler, Koenig, Lockwood, Oberhofer, Schroeck

ISBN-13: 978-0-13-703571-7

Architecture for the Intelligent Enterprise: Powerful New Ways to Maximize the Real-time Value of Information

In this book, a team of IBM's leading information management experts guide you on a journey that will take you from where you are today toward becoming an "Intelligent Enterprise."

Drawing on their extensive experience working with enterprise clients, the authors present a new, information-centric approach to architecture and powerful new models that will benefit any organization. Using these strategies and models, companies can systematically unlock the business value of information by delivering actionable, real-time information in context to enable better decision-making throughout the enterprise–from the "shop floor" to the "top floor."



Dynamic SOA and BPM

Best Practices for Business Process Management and SOA Agility

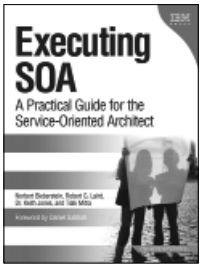
By Marc Fiammante

ISBN-13: 978-0-13-701891-8

Achieve Breakthrough Business Flexibility and Agility by Integrating SOA and BPM

Thousands of enterprises have adopted Service Oriented Architecture (SOA) based on its promise to help them respond more rapidly to changing business requirements by composing new solutions from existing business services. To deliver on this promise, however, companies need to integrate solid but flexible Business Process Management (BPM) plans into their SOA initiatives. ***Dynamic SOA and BPM*** offers a pragmatic, efficient approach for doing so.

Related Books of Interest



Executing SOA

A Practical Guide for the Service-Oriented Architect

By Bieberstein, Laird, Jones, Mitra

ISBN-13: 978-0-13-235374-8

This book follows up where the authors' best-selling Service-Oriented Architecture Compass left off, showing how to overcome key obstacles to successful SOA implementation and identifying best practices for all facets of execution—technical, organizational, and human. Among the issues it addresses: introducing a services discipline that supports collaboration and information process sharing; integrating services with preexisting technology assets and strategies; choosing the right roles for new tools; shifting culture, governance, and architecture; and bringing greater agility to the entire organizational lifecycle, not just isolated projects. Executing SOA is an indispensable resource for every enterprise architect, technical manager, and IT leader tasked with driving value from SOA in complex environments.

A Practical Guide to Distributed Scrum

Woodward, Surdeck, Ganis

ISBN-13: 978-0-13-704113-8

The Business of IT

How to Improve Service and Lower Costs

Ryan, Raducha-Grace

ISBN-13: 978-0-13-700061-6

Viral Data in SOA

An Enterprise Pandemic

Fishman

ISBN-13: 978-0-13-700180-4

Disciplined Agile Delivery

A Practitioner's Guide to Agile Software Delivery in the Enterprise

Ambler, Lines

ISBN-13: 978-0-13-281013-5

Outside-in Software Development

Kessler, Sweitzer

ISBN-13: 978-0-13-157551-6

This page intentionally left blank

Patterns of Information Management

This page intentionally left blank

IBM PRESS

Patterns of Information Management

Mandy Chessell
Harald C. Smith

IBM Press
Pearson plc

Upper Saddle River, NJ • Boston • Indianapolis • San Francisco
New York • Toronto • Montreal • London • Munich • Paris • Madrid
Cape Town • Sydney • Tokyo • Singapore • Mexico City

ibmpressbooks.com

IBM Press Program Managers: Steven M. Stansel, Ellice Uffer

Cover design: IBM Corporation

Editor In Chief: Bernard Goodwin

Marketing Manager: Stephane Nakib

Publicist: Heather Fox

Editorial Assistant: Michelle Housley

Managing Editor: Kristy Hart

Designer: Alan Clements

Project Editor: Andy Beaster

Copy Editor: Karen Annett

Indexer: Lisa Stumpf

Compositor: Gloria Schurick

Proofreader: Sarah Kearns

Manufacturing Buyer: Dan Uhrig

Published by Pearson plc

Publishing as IBM Press

IBM Press offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales, which may include electronic versions and/or custom covers and content particular to your business, training goals, marketing focus, and branding interests. For more information, please contact

U.S. Corporate and Government Sales
1-800-382-3419
corpsales@pearsontechgroup.com.

For sales outside the U.S., please contact

International Sales
international@pearson.com.

Library of Congress Cataloging-in-Publication Data is on file.

ISBN-13: 978-0-13-315550-1
ISBN-10: 0-13-315550-1

Text printed in the United States on recycled paper at Courier in Westford, MA.
First printing: May 2013 with corrections September 2013

Some day, on the corporate balance sheet, there will be an entry which reads, "Information"; for in most cases, the information is more valuable than the hardware which processes it.

—Rear Admiral Grace Hopper

To my family and friends who have supported and encouraged me over the years, with love:
Sarah, Jane, Ray, Terri, Ben, Ella, Ian, Kay, Keith, Dan, and Chris.

—Mandy Chessell

To my family: Mary, Rebecca, Cecilia, and Paul
and my parents: Clyde and Ellen
For their love and support now and over the years.

—Harald Smith

Contents at a Glance

This page intentionally left blank

Table of Contents

Foreword by Rob High

In the early history of humans—when we were barely distinguishable from other animals—we protected ourselves with the construction of shelters made from branches, leaves, and mud. To suggest that these structures had architecture is being generous. But to the extent that their architectural style could be discerned at all, at best you could argue their architecture was forged from necessity and practicality. Dried mud helped keep the leaves secured and leakproof. Branches provide strength across broad spans. Leaves knitted everything together. Eons of experience and mistakes drove practices—often varying substantially from one region to another based on the available materials and climate conditions of each region. Form strictly followed function. No one, at that time, ever stopped to draw out their shelter design—critiquing it for its style and propitiousness.

Likewise, the early years of information systems implementation were driven out of necessity—leveraging the materials and practices that were available to us. Barely discernable architectures were more derived than prescribed. But, as with our ancestral laborers, we found the need to build more sophisticated and ever-larger answers to our problems. And, as before, we discovered that with a little planning, engineering, and standardized construction techniques, we could handle the task more easily. We could adapt to new functions, and we could apply form and aesthetics to the things we constructed. The results are not only more pleasing, but they are more reliable, more efficient, more economical, and more adaptable to changing conditions. There is not only derived architecture, but we can employ architecture to coerce solutions to better meet our needs.

At the heart of architecture are patterns that shape and style the materials we work with and the techniques that inform the practices of pattern adoption. Patterns form the building blocks of construction. Techniques tell us how best to select and assemble those building blocks to achieve the results we are seeking.

The idea of using patterns and techniques in the construction of information management is exactly what this book is about. Information patterns form the basis of a whole new architectural approach to systems design that, like its analog in the construction industry, is essential to assuring durability, usability, flexibility, and utility of IT solutions. The patterns presented here

are not hypothetical, but rather have been forged from decades of experience in the field. They capture the bittersweet results of literally thousands of person-years of effort—trying different approaches, abandoning the ones that failed, refining the ones that showed promise, latching on to the ones that proved to work well, and promoting the ones that exceeded expectations.

The language of information patterns is a formalism of expression. It standardizes an approach to representing ideas that enables clarity and precision of communication. It allows us to exchange a common understanding of those ideas and to manipulate those ideas to create even bigger, more profound ideas. And although the basic ideas captured in the language are not new, the ability to express those ideas in a formal language is both novel and profoundly innovative. It makes it possible to unleash the value of information patterns, and build even greater value more effectively.

As the former chief architect of IBM®'s SOA Foundation, I have long advocated the use of architecture for ensuring the utility of information systems solutions. These information patterns build on the traditions of Service Oriented Architecture (SOA) and, more important, fill a critical gap in SOA by addressing the correlation between services and information in our IT solutions. In keeping with the SOA tradition, these patterns are not just the raw materials of construction for IT developers, but are essential to aligning the objectives of IT and business—driving the form of the system to both enable the business to respond quickly to changing business conditions and to even compel the business to motivate changes that will gain it a competitive advantage in its marketplace.

The use of information patterns is imperative for modern information system design. This book is a must-read and I strongly encourage you to apply these techniques in your practices.

Rob High
IBM Fellow
CTO, IBM Watson™
Former chief architect, SOA Foundation

Preface

About This Book

Information is the heart of any organization's operation. It defines who is involved, what activities are taking place, and the assets being brought to bear to create its goods and services. Managing information takes a multidisciplinary approach because it pervades every aspect of the organization's life. We must consider where information comes from; consider how it is distributed, protected, governed, and monitored; as well as ensure it is used appropriately at its destination.

From a technical point of view, a myriad of technologies have emerged to tackle different aspects of information management. There is the Service Oriented Architecture (SOA) technology that makes information available to remote systems; databases and files systems to store information; messaging technology that sends notifications between systems; extract, transform, load (ETL) technology for moving and transforming large quantities of information; federated queries; big data technologies; replication technology; data quality and metadata tools; distributed security technology; analytics; archiving processes; and many more. For the architect, the choice of technology is overwhelming and there is a strong temptation to just stick with the technology we know well. The result is that our designs can become unbalanced, with too much focus on one dimension of the problem.

This book contains architecture patterns that characterize the typical information issues associated with distributed systems. They demonstrate how the seemingly competing technologies for information management, SOA, and business process management can be blended to create an effective, interconnected, and ordered IT landscape, making it both manageable and efficient.

The patterns are built around a supply chain metaphor. Information is supplied to the organization, processed, moved around, processed some more, and then output as some form of information product, such as a report or dashboard. Through this analogy, we cover how an organization can manage its information to support its operation effectively, balancing quality, availability, breadth, precision, and timeliness with cost.

Information is not a physical asset, of course, so our notion of an "information supply chain" must extend the physical manufacturing and distribution supply chain metaphor with the recognition that information rarely moves from point to point as a discrete, whole unit. It is constantly being copied, transformed, renamed, partitioned, merged, updated, and deleted. More subtly, there are multiple versions of the "truth." People's interpretation and assumptions around the information they work with are highly contextual, and so when information is shared, it must be transformed to match the expectations and context of the new consumers before they will trust it to support their business.

Patterns of Information Management explains how information is transformed, enriched, reconciled, and redistributed along the information supply chain. The aim is to shape the way systems are integrated to create an orderly flow of information that can be reused and synchronized at key points in the processing.

A pattern-based approach is powerful because the resulting system behavior is determined by the way technologies are combined; the design choices are heavily affected by nonfunctional requirements, such as the amount of information, the arrival rates of new values, the level of quality that can be assumed, and the processing required to make the information useful. The patterns link together into a pattern language that spans from the holistic system level views for enterprise architects down to the design patterns of integration developers. The resulting pattern language enables the architect to make reasoned decisions about the applicability of each alternative approach and the inherent consequences of the choice.

Intended Audience

Patterns of Information Management is intended for enterprise architects, information architects, and solution architects who are responsible for defining how information systems should be linked together in order to synchronize, manage, and share information. Students and practitioners alike will find that the patterns create a framework in which to organize their existing experience and broaden their knowledge.

How to Use This Book

This book is principally a reference book. Chapters 1 and 2, "Introduction" and "The MCHS Trading Case Study," respectively, set the scene, explaining how the pattern language is structured and walking through some examples. It is recommended that you read both of these chapters before using the patterns.

The remaining chapters contain the pattern definitions themselves. These chapters can be used to deepen your knowledge of the topic or as a pattern reference during a particular project. Because there are so many patterns, it is not possible, or necessary, to learn all of the patterns before starting to use them in a project. The patterns are organized into small, related groups that focus on a particular aspect of information management. We suggest that you start with a pattern

group that is particularly relevant to your work. Study the pattern descriptions, paying attention to the differentiating aspects of each pattern, and map them to your project.

The content can be used in various ways:

- The pattern names define a vocabulary to discuss issues and technology options related to information architecture, governance, and management. The pattern names become the nouns and verbs of your design discussions, enabling you to characterize and choose between the options available.
- The pattern icons can be used in whiteboarding sessions and for documenting design decisions in reference architectures and solution specifications.
- The pattern descriptions provide design guidance for specific information supply chains and solutions.
- The pattern language provides a foundation for setting architectural standards and reference architectures. An organization can select the patterns it wants to support and develop reference implementations for them.
- The pattern language provides material for education and training in information architecture.

As you become familiar with one group of patterns, you can turn your attention to another—iteratively growing your knowledge. The consistency built into the pattern language will accelerate your learning process and very quickly you will become proficient in a significant working set of patterns.

Structure of the Book

Patterns of Information Management is divided into the following chapters:

Chapter 1, "Introduction"

Chapter 1 is an introduction to the topic of information management and the challenges that architects face. It explains what a pattern is, the conventions we use in documenting a pattern, and how patterns are linked together to form a pattern language.

Chapter 2, "The MCHS Trading Case Study"

Chapter 2 uses a case study involving a fictitious company called MCHS Trading to illustrate how the patterns of information management can be used in a project setting. These projects cover defining an information strategy; introducing a data warehouse and management reporting; introducing master data management hubs for operational master data; improving how information is governed, monitored, and managed; exchanging information with external parties; and using predictive analytics.

Chapter 3, "People and Organizations"

Chapter 3 is the first of the pattern definition chapters. It covers patterns relating to the way an organization is operating because this has a major impact on the way that information is managed. It describes the following pattern groups:

- Information centric organizations
- Information users

These patterns help in the definition of your information strategy and information governance program.

Chapter 4, "Information Architecture"

Chapter 4 covers the nature of information and how to think about its structure, meaning, and organization. This includes patterns that describe the different life-cycle patterns of information, models, metadata, information supply chains, and the solutions you may implement to improve the management of information. It includes the following pattern groups:

- Information elements
- Information identification
- Information provisioning
- Information supply chains

These patterns help you understand and plan the overall flow of information through your organization's systems.

Chapter 5, "Information at Rest"

Chapter 5 contains pattern descriptions describing how information is accessed and stored. This includes descriptions of the different types of servers that support information management. The pattern groups in this chapter are as follows:

- Information services
- Information collections
- Information entries
- Information nodes

These patterns help to shape your thinking when deploying new capability into the IT environment, such as a new application, master data management hub, data warehouse, analytics, or reporting platform.

Chapter 6, "Information in Motion"

Chapter 6 documents the patterns for moving information between servers, including services, information routing, filtering, consolidation, and distribution. It contains the following pattern groups:

- Information requests
- Information flows

These patterns focus on how information is moved between the information nodes to satisfy the needs of their information processes.

Chapter 7, "Information Processing"

Chapter 7 describes the different types of processing that is performed on information and how it is triggered. The processes include business, movement, transformation, quality management, search, and analytics processes. It contains these pattern groups:

- Information triggers
- Information processes

These patterns help you categorize and select how information should be processed and maintained.

Chapter 8, "Information Protection"

Chapter 8 covers the patterns for protecting information, including validating, transforming, enriching, and correcting information; security; and monitoring. It includes three pattern groups:

- Information reengineering steps
- Information guards
- Information probes

Once you have your information supply chains mapped out, these patterns help you design how the information can be protected end to end.

Chapter 9, "Solutions for Information Management"

Chapter 9 covers the information solution pattern group. These patterns describe a selection of information projects that improve the information management of an organization's information systems. Solutions incorporate a set of patterns focused on addressing particular information problems, though always in context with existing information landscapes so they must integrate with, enhance, or incorporate existing systems and information.

Appendices

The appendices include a glossary of technical terms and technology types that are mapped to the relevant patterns.

What Is Not Covered

This is a design pattern language and so does not recommend specific technologies and products. It aims to describe classes of technology, where they should be used, and the consequences of using them. This book is also intended to be an overview of the information management landscape. It does not include implementation details, particularly where there is plenty of literature covering these details. Finally, there are references to information governance in the patterns. Information governance is a supporting discipline for information management that involves business controls, procedures, and related technology. It is beyond the scope of this book to cover these aspects properly.

Further Information

There is an IBM developerWorks® community for the *Patterns of Information Management* where additional details, examples, and discussions are being posted. To join, visit:
https://www.ibm.com/developerworks/community/groups/community/PoIM

This community provides you with an opportunity to ask questions about the patterns and suggest improvements.

Acknowledgments

The pattern language has been through many iterations, discussions, and workshops to get it into a state that is suitable for publishing. We would like to thank the following people who have provided help and feedback during the process.

From the EuroPLoP 2011 Pattern Writers Conference (where the **I**NFORMATION **N**ODE patterns were reviewed), there is our shepherd, Tim Wellhausen, and the members of the writer's workshop: Klaus Marquardt, Markus Gaertner, Andreas Fiesser, and Joerg Pechau.

IBM has provided us with an envirionment where we have been able to develop and share the material with experts from many companies, industries, and software disciplines, drawing on their years of experience delivering complex projects and systems. As a result, we received valuable feedback from many subject matter experts, including Dan Wolfson, Dougal Watt, Steve Lockwood, Kyle Brown, Loretta Mahon Smith, Mike McRoberts, Rob High, Martin Wildberger, Richard Hedges, Stephanie Hazlewood, Patrick Dantressangle, Mukesh Mohania, Ivan Milman, Lena Woolf, David Radley, Larry Yusuf, Susan Visser, Sarah Tee, Herb Berger, Prasad Vempati, Keith Enhagen, Eleomara Goncalves, and Rickey Tang.

We would also like to thank the members of Pearsons plc who have expertly guided and supported us in our first publishing effort—in particular, Bernard Goodwin and Michelle Housley.

About the Authors

Mandy Chessell

FREng CEng FBCS

Mandy has worked for IBM since 1987. She is an IBM Distinguished Engineer, IBM Master Inventor, and member of the IBM Academy of Technology Leadership Team. As the chief architect for InfoSphere® Solutions in IBM's Software Group, Mandy designs common information integration patterns for different industries and solutions.

In earlier roles, Mandy's work has focused on transaction processing, event management, business process management, information management, and model-driven development. This breadth is reflected in her invention portfolio, which to date stands at over 50 issued patents worldwide.

Outside of IBM, Mandy is a fellow of the Royal Academy of Engineering and a visiting professor at the University of Sheffield, UK. In 2001, she was the first woman to be awarded a Silver Medal by the Royal Academy of Engineering, and in 2000, she was one of the "TR100" young innovators identified by MIT's *Technology Review* magazine. In 2006, she won a British Female Innovators and Inventors Network (BFIIN) "Building Capability" award for her work developing innovative people and the BlackBerry "2006 Best Woman in Technology - Corporate Sector" award. More recently, she was granted an honorary fellowship of the Institution for Engineering Designers (IED) and she won the "2012 everywoman Innovator of the Year."

For more information on Mandy's publications, see http://en.wikipedia.org/wiki/Amanda_Chessell.

Harald Smith

Harald has worked for IBM since 2005. Harald is a software architect in IBM's Software Group specializing in information quality, integration, and governance products, and is IBM certified in delivering IBM Information Management solutions. In this role, he develops best practices, methodology, and accelerators for common information integration use cases.

Harald has 30 years of experience working with data quality products and solutions; product and project management; application development and delivery; system auditing; technical services; and business processes across the software, financial services, healthcare, and education sectors. Harald was the product manager at Ascential® Software and IBM responsible for designing and bringing the IBM InfoSphere Information Analyzer product to market as a key component in IBM's information quality portfolio. He has been issued three patents in the field of information quality and rule discovery and was recently recognized as an IBM developerWorks Contributing Author.

His publications include the IBM developerWorks articles "The information perspective of SOA design" [parts 6, 7, and 8], "Use IBM WebSphere® AuditStage in a federated database environment," "Using pre-built rule definitions with IBM InfoSphere Information Analyzer,"

"Designing an integration landscape with IBM InfoSphere Foundation Tools and Information Server" [part 1], and "Best practices for IBM InfoSphere Blueprint Director" [parts 1 and 2]. For the IBM InfoSphere Information Server documentation, Harald contributed to the "IBM InfoSphere Information Analyzer Methodology and Best Practices Guide" and "IBM InfoSphere Information Server Integration Scenario Guide"; he has also contributed to three IBM Redbooks®.

CHAPTER 1

Introduction

In software engineering, we are taught that abstraction, modularity, and information hiding are useful approaches when breaking down a set of requirements into manageable chunks for implementation. This book makes use of these principles to tackle the synchronization of information between IT systems.

Islands of Information

Most organizations use specialized IT systems called applications to run their operations. Each application supports a particular aspect of the business, either for the whole organization or a group within it. There may be applications for order taking, for billing, for distribution of goods, for management of employee data, and many more.

An application will store the information it is processing in a persistent store for later reference. This information store is often a private resource for the application. Over time, this store contains important details about the people with whom the organization is interacting, what assets they use, how, when, and why.

A healthy organization will develop and grow—and this change drives changes into its applications, affecting their function and scope. It can also lead to duplication of function:

- When two organizations merge, they can end up with at least two applications for each function.
- When a new product line or channel to market is introduced, an organization may choose to introduce a new application to support it, to avoid the possibility of disrupting the established business or to implement it faster.
- Multinational organizations find they need separate applications for different countries, or trading regions, to handle local customs and regulations.

Careful management and constant rationalization may reduce the number of applications so there is little or no overlap in function. However, an application is a complex mix of software and hardware. It takes considerable engineering effort to develop it, and so once the investment is made, an application is expected to have a long life (5–15 years). Ripping it out and replacing it can be expensive and difficult and so an organization may choose to maintain multiple applications for the same function.

When there are two applications covering the same function, information about that function is split between the two applications and is typically stored in a different format. Even when all applications support unique functions, there is still an overlap in the information that they hold. This is the information that describes the core interests of the organization, such as customers, suppliers, products, contracts, payments, assets, employees, and many more.

Over time, the private information stores of an organization's applications become islands of duplicated and inconsistent information. This affects the efficiency of an organization and its ability to operate in a cost-effective, flexible, and coherent unit.

This book seeks to address the challenge of effective information management. How does an organization improve its management of information, working with the applications it already operates, to ensure it knows what its assets are, what it is working on, what commitments it has agreed to, how well it is performing, and how it can improve its operation?

Introducing MCHS Trading

MCHS trading is a fictitious trading company used in this book to illustrate different approaches to information management. MCHS Trading sells goods through four channels: on the Internet, via mail order, via a call center, and through physical shops. Due to differences in requirements, orders are taken by three different applications: E-shop for the Internet orders, Mail-shop for mail order and the call center, and the Stores application to support the needs of the physical stores. These applications were introduced incrementally as MCHS Trading opened the new channels to its customers.

Orders are fulfilled and money is collected through the Shipping and Invoicing applications, respectively. The E-shop, Mail-shop, and Stores applications send the order details to the Shipping application, which in turn forwards them on to the Invoicing applications as soon as an order is dispatched. This is illustrated in Figure 1.1.

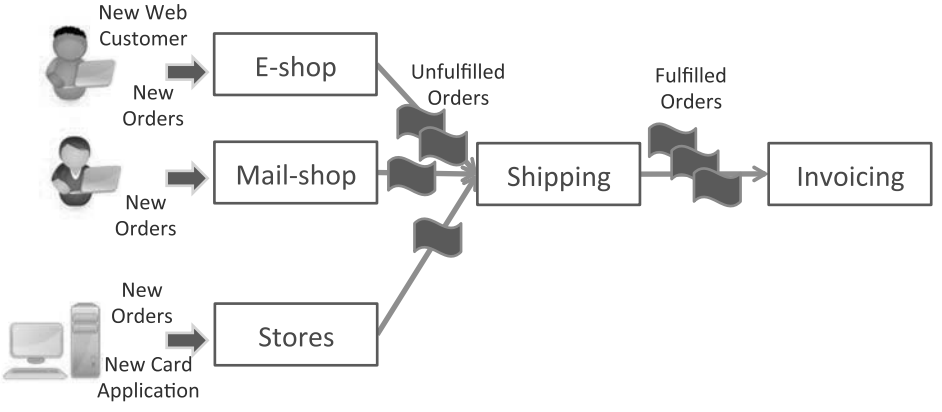


Figure 1.1 MCHS Trading's order-processing systems.

From a functional point of view, this is a rational separation of concerns. Each application has a clearly delineated set of responsibilities and the process works—customers get the goods they ordered and the correct money is collected in exchange for the goods.

However, when the information stores are added to the picture (see Figure 1.2), you can see that the details about customers, products, and orders are replicated across the systems. Why? Because this information is core to MCHS Trading's business and so every application needs it in some form or another.

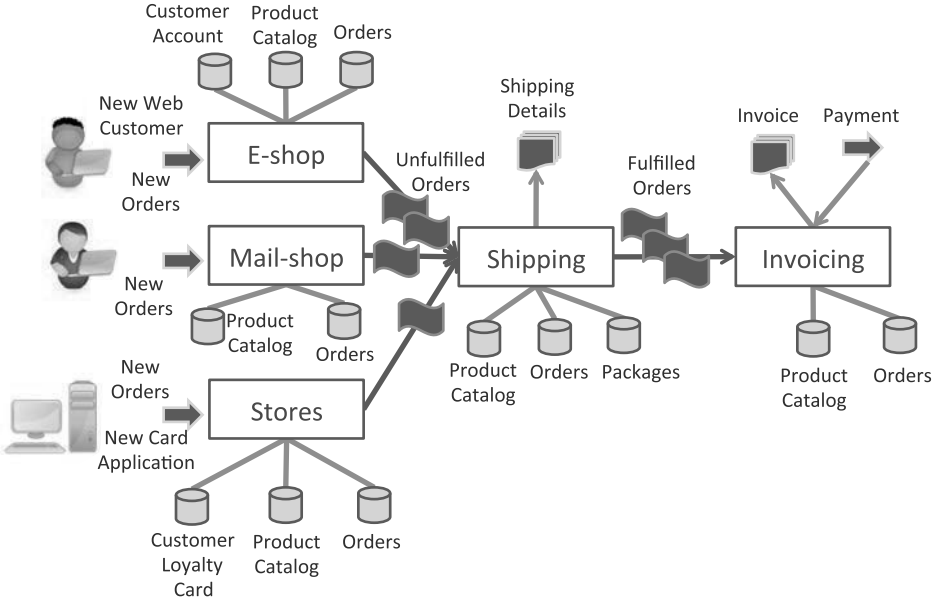


Figure 1.2 Information stores supporting order processing.

Failing to synchronize this information effectively leads to inflexibility and inefficiencies in the organization that can have an impact both internally and externally.

Consider an individual customer, Alistair Steiff. He has registered for a loyalty card, which is handled by the Stores application, and he also uses the E-shop and the mail order channel from time to time. The E-shop keeps a record of Alistair in the form of a customer account. This is different from the loyalty card account. The Mail-shop application takes Alistair's details with each order. It has no capability to maintain his details for the next time he orders something through that channel, which Alistair finds a little annoying.

Alistair experienced another issue when he moved to a new address. Although he updated his address in his E-shop account, MCHS Trading kept sending his loyalty card statements to his old address. He tried phoning MCHS Trading's call center but they could not help because they

were only set up to take new orders. He had to write a letter to MCHS Trading to get the loyalty card address updated.

Within MCHS Trading, there is also frustration with the current systems. It is difficult to understand the buying patterns of its customers:

- To understand which new products would be of interest to an individual
- To understand how an individual interacts across each of MCHS Trading's sales channels

There are two issues here. First, the applications only store information that is relevant to their operation—so it is hard to see the complete picture when the details are spread among the applications. Second, applications are designed to reflect the current operational state of their work. They may keep historical data, but not in the form that is conducive to analysis of trends and anomalies.

This lack of insight is inhibiting MCHS Trading's ability to grow and there is a need to introduce new management reporting capability to understand how the business is really performing.

Improving an Organization's Information Management

The problems that MCHS Trading is experiencing are typical for many organizations. Information is distributed across multiple applications for operational efficiency, leaving the information duplicated, fragmented, and often inconsistent. As a result, the organization cannot act in a coordinated manner:

- They find it hard to get an overall picture of how well the organization is performing. This requires a consolidated view of their business activity, showing the current position, along with a historical perspective for comparison. For example, MCHS Trading would want to understand how many customers it has, what types of products its customers are interested in buying, how this is changing over time, how efficient the delivery process is, who its best suppliers are, and much more if it is to maintain its market leadership position.
- The quality of information varies from application to application. This means different parts of the organization are operating on different facts that could lead to different decisions being made for the same situation.
- The internal fragmentation of the information is often exposed outside of the organization, creating poor customer service or missed opportunities. This was Alistair Steiff's experience when he tried to change his address—he had to ensure it was updated in each application—using a different process for each one.

- External regulators are skeptical that the organization's reported results are accurate when ad hoc processes are used to create them because it is difficult to explain where the information came from and how the results were calculated.
- Information is not retained for the required amount of time—or too much information is retained for too long. Either case can inhibit the ability to find the right information in a timely manner.
- Failures in the mechanisms that move information around can corrupt the integrity of the organization's information. These failures may be errors in the information itself, which means it cannot be transferred, or errors in the implementation of the mechanism, or a failure to initiate some processing in time, resulting in missing information. These failures may not be detected for some time and can be extremely difficult to resolve.

An organization needs to maintain a strong core of information to run the business. This requires a focus on how key information is created, processed, and stored within its IT systems:

- They must *optimize where information is located relative to the workload* that is using it—ensuring copies are taken in a thoughtful way and these copies are supported with mechanisms to maintain or remove them as new information becomes available.
- *Related information should be correlated together* to create a complete picture of the organization's activities.
- *Obsolete information must be removed* to save storage and reduce the processing effort. Vital historical information needs to be retained.
- *Information must be protected* from inappropriate use, restored after a failure, and, despite the fact that most organizations have their silos and cliques, *the right information needs to be exchanged and presented to the right people at the right time*.

This is hard to achieve. Technology is often focused on providing function to the business rather than managing information. Specialist information management technology helps, but it has to be blended with the existing infrastructure. The blending process creates emergent properties.

Emergent properties are those characteristics that "appear" when components are combined together in a particular configuration—rather than being inherent properties of any one of the components. This is similar to the behavior of colors when you mix them together. If you combine blue and yellow, you get the color green. Green is an emergent property because it is not present in either the blue or the yellow—and only emerges when they are combined.

When we combine technology together, we also get emergent properties. These emergent properties may be additive, or they may override some of the components' original capabilities.

Consider two applications sited on opposite coasts of the United States. The application on the East Coast needs to regularly send information to the West Coast application. However due to the different time zones in which they operate, there is a 3-hour period of its operation when the

West Coast application is not available. It is not possible to extend the period of operation of the West Coast application—and the East Coast application is not capable of buffering the information until the West Coast application comes online.

The solution is to provide a new database that is available whenever either application is online (see Figure 1.3). The East Coast application writes information into the database. The West Coast application processes the information in the database when it becomes available.

This is a common integration approach. The database in the middle is acting as a staging area. Its effect in the integration is to expand the time window that information can be transferred between the two applications. It may also improve the resiliency of the integration because the East Coast application is no longer affected by the occasional outages of the West Coast application. However, the downside is that when both applications are available, the staging area slows down the transfer of information between the two applications because there is a small delay between the East Coast application writing the information and the West Coast application picking it up. The time it takes to transfer information between two systems is called the latency of the information transfer.

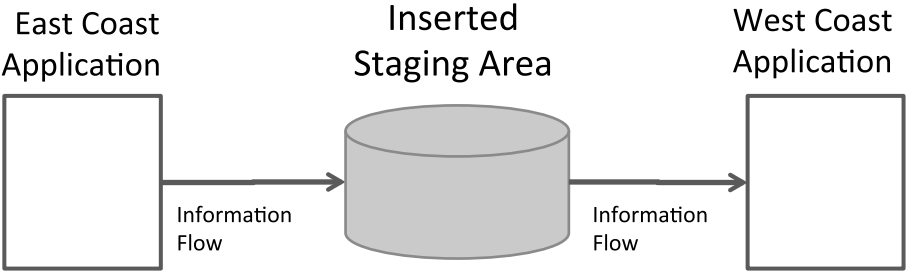


Figure 1.3 Using a staging area between two applications.

The increased availability, latency, and resilience are all emergent properties of the integration. In general, the emergent properties relate to non-functional characteristics that may not be evident until the integration is in production. This is why architects like to use tried and tested approaches where the emergent properties are well understood.

This book describes how the flow of information between applications and other systems should be designed, calling out the emergent properties as they occur.

The aim is to shape the placement of workload and information stores within the IT systems to create an orderly flow of information that guarantees the quality of the results. The material is presented as a set of connected software design patterns, called a pattern language.

Patterns and Pattern Languages

A software design pattern defines a proven approach to solving a problem. The solution described in the pattern is typically a set of components that are interacting in a particular configuration. The pattern explains why this approach works, its associated trade-offs, and resulting benefits and liabilities. It also links to other patterns that may:

- Provide an alternative approach
- Provide a complementary capability
- Describe an approach to implementing a component that is named in the pattern's solution

This linking together of related patterns creates what we call a pattern language.

Every pattern in this pattern language has a name that summarizes the solution it represents in a succinct manner. For example, there is a pattern called **I**NFORMATION **C**OLLECTION that describes a collection of related information. Notice that the pattern name is written using the small capitals formatting. This formatting convention is used wherever the pattern is first referenced in a section. The name of a pattern can act as shorthand for the solution during design discussions.

Choosing the terminology for the pattern names has been a challenge because the pattern language covers multiple architectural disciplines. Where possible, we have used industry standard names for the concepts and components exposed in the pattern language. However, we have found it necessary to introduce new terminology whenever there is conflicting nomenclature, or no obvious name exists.

Every pattern has an icon that can be used as a visual reminder of the pattern, particularly when whiteboarding and documenting solutions. The information collection icon is shown in Figure 1.4.

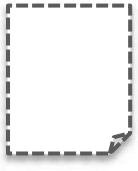


Figure 1.4 Icon for the information collection pattern.

The patterns are built on a common component model. This means that a pattern can be used as a component in the solution described by another pattern. When this occurs, the icon of the pattern is used in the solution diagram of the consuming pattern. For example, Figure 1.5

shows the solution for the staging area introduced in the previous section. It is built from a variety of pattern icons.

The meaning of the icons in the diagram, and the details of the patterns behind them, will become familiar to you as you work with the pattern language. The purpose here is to illustrate how the icon of one pattern, the information collection in this case, can be used in the solutions of other patterns.

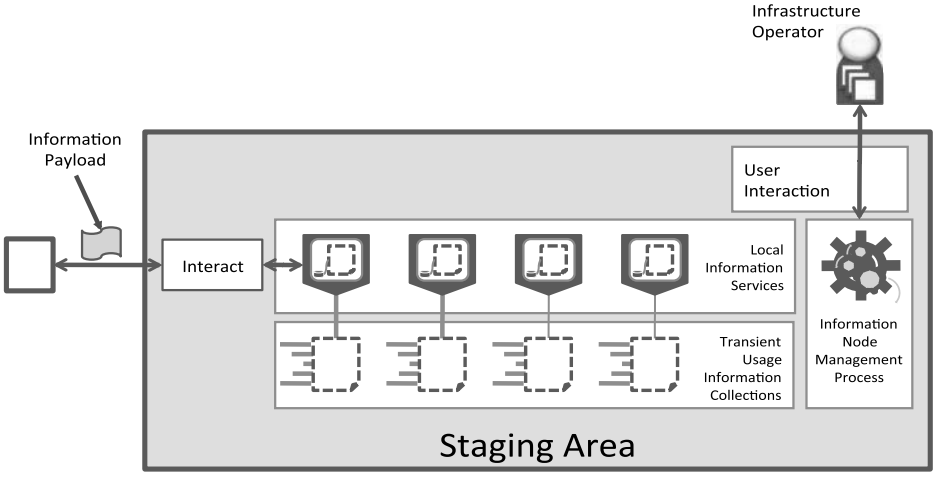


Figure 1.5 The solution diagram for a staging area showing the use of the information collection icon.

Each of the patterns of information management can be used independently. However, the real value of a pattern language is the ability to compare and contrast different approaches to resolving a situation.

The patterns of information management that are relevant to a particular situation are collected together in a pattern group. Each pattern group has a lead pattern that describes the core principles and capabilities of the group. The pattern group is named after the lead pattern. The other patterns in the pattern group enhance one or more characteristics of the lead pattern to support a more specialized situation.

You may have noticed that in Figure 1.5, the information collection icon appears slightly modified in the solution diagram with five lines coming out of the left side (as shown in Figure 1.6). The modification to the information collection icon denotes that the staging area uses specialized information collections described by the **T**RANSIENT **S**COPE pattern.

The modified icon visually represents that there is a relationship between the information collection and transient usage patterns. They are, in fact, from the same pattern group along with other patterns called **L**OCAL **S**COPE and **C**OMPLETE **S**COPE.

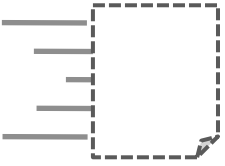


Figure 1.6 Icon for the transient scope pattern.

To make it easy to compare and contrast the patterns in a group, each group of patterns begins with a table of pattern summaries, called patlets. A patlet shows a pattern's icon, name, short problem statement, and summary of the solution. The aim of the patlets is to help you quickly discover and navigate to the pattern you need. Table 1.1 shows the patlet for the **I**NFORMATION **C**OLLECTION pattern, which is the lead pattern in its group.

Table 1.1 Sample Pattern Summary for Information Collection

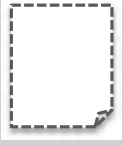
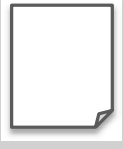
Icon	Pattern Name	Problem	Solution
	**I**NFORMATION **C**OLLECTION	Information must be organized so it can be located, accessed, protected, and maintained at a level that is consistent with its value to the organization.	Group related information together into a logical collection and implement information services to access and maintain this information.

Table 1.2 shows the related scope patterns from the same group. Notice that the icons are all variations of the information collection icon.

Table 1.2 Patlet Table for the Scope Patterns in the Information Collection Group

Icon	Pattern Name	Problem	Solution
	**C**OMPLETE **S**COPE	An information process needs to perform an activity once for each instance of a particular subject area (such as a customer, product, order, invoice, shipment, etc.) that occurs within the information supply chain.	The information process needs to use an information collection that stores a single information entry for each instance of the subject area that occurs within the information supply chain. Such an information collection is said to have a complete scope.

Icon	Pattern Name	Problem	Solution
	**L**OCAL **S**COPE	The implementations of the information processes hosted within an information node assume they are in complete control of changes to the information they use.	Provide information collections within the information node for the sole use of its information processes. These information collections will then only have information entries that are created by the locally hosted information processes. These types of information collections are said to have a local scope.
	**T**RANSIENT **S**COPE	An information node needs to provide temporary storage for information entries that are being continuously added and removed by the information processes.	Create an information collection to temporarily store the information entries in the information node. From time to time, the information entries stored in this information collection will change, and so we say this collection has transient scope.

The detailed pattern descriptions follow the patlet tables. Many different styles and heading structures have been successfully used to describe software patterns. We have chosen to use one of the formats recommended by The Open Group Architecture Framework (TOGAF®)[1] with the following subsections:

- **Context—**The situation where it is appropriate to consider using the pattern
- **Problem—**A description of the problem that this pattern solves
- **Example—**An example of the problem
- **Forces—**The factors that make this problem hard to solve
- **Solution—**A description of solution components and how they are assembled together
- **Consequences—**The benefits and liabilities of using the solution
- **Example Resolved—**How the example described is resolved using the pattern
- **Known Uses—**References to well-known technologies and approaches that support the pattern
- **Related Patterns—**Links to other relevant patterns in the pattern language

1. http://www.togaf.info

These sections provide a way to bring together a variety of information into a well-formed structure that summarizes the essence of the pattern. Together, they enable you to make reasoned choices of approach for the solution you are building.

Basic Components in the Pattern Language

A good place to start learning about the *Patterns of Information Management* is a pattern called **I**NFORMATION **P**ROVISIONING. This describes how five basic components interact to receive, process, and produce information.

Figure 1.7 comes from the information provisioning pattern description and shows the components and how they interact. All of these relationships are many-to-many.

Notice that the information collection pattern introduced previously is shown as a component in the information provisioning pattern along with some additional patterns: **I**NFORMATION **U**SER, **I**NFORMATION **T**RIGGER, **I**NFORMATION **P**ROCESS, and **I**NFORMATION **S**ERVICE.

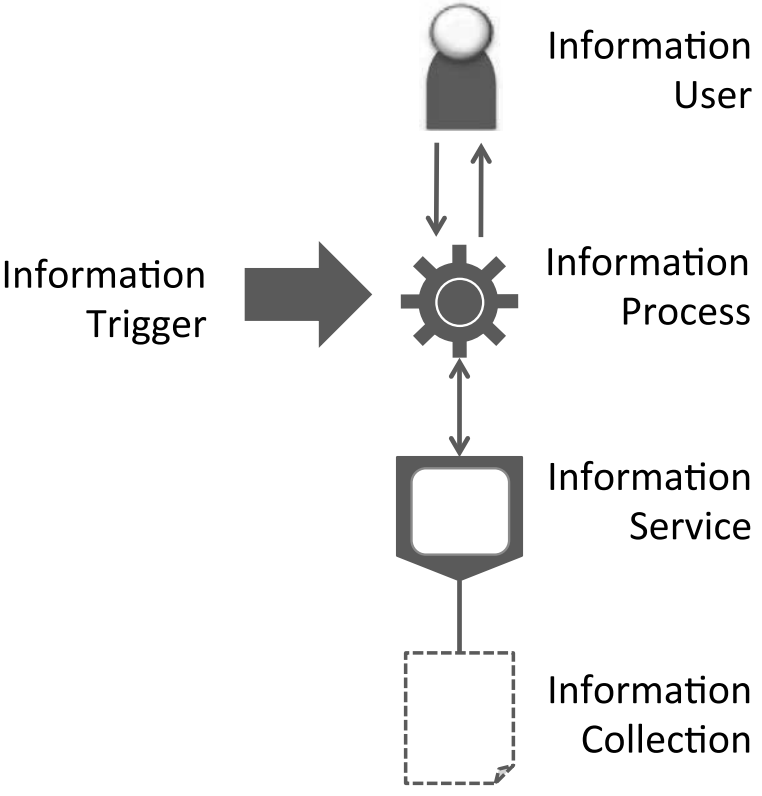


Figure 1.7 Five basic components for processing information.

The information user is a person working with the organization's information. This person may be an employee or someone outside the organization such as a customer or a supplier. Each kind of information user has his or her own requirements for the kinds of information needed, where, and when. The information user is both a consumer of existing information and a contributor of new information. The information user works with user interfaces that are controlled by information processes.

The information processes perform the automated processing of the organization. There are many different kinds of information processes—but, collectively, they are the mechanisms by which information is received, transformed, and produced in some form or another.

An information trigger starts an information process. This may be the result of an information user request, a scheduler, an event being detected, or the need for more information. The information trigger passes the information process some context information that describes why it is being started. The information process augments this context information with information from the information users and other information known to the organization.

An information process accesses any additional information it needs through well-defined interfaces called information services. An information service provides a specialized view of the information that the information process needs. It is able to locate the requested information from a variety of sources and transform it into a format suitable for the requesting information process. An information service retrieves stored information from information collections.

The information collection manages a collection of related information. Typically, the contents of an information collection relate to the same subject area. However, an information collection may contain information that was collected from the same activity, such as the results of an experiment, from the same source, such as readings from a sensor, or from the same time period, such as social media extracted for a specific period of time.

The information provisioning pattern is a lead pattern of a pattern group. It recognizes that the information a person sees through a user interface is different from the way it is stored and shows the layering of components used to manage the mapping. The rest of the pattern group describes how information is provisioned when multiple systems are involved.

Information Integration and Distribution

As already discussed, an organization will have many applications and other kinds of IT systems. Each of these systems will host their own information processes, information services, and information collections.

We use the **I**NFORMATION **N**ODE pattern to represent the general concept of a system. You may want wish to think of this as a physical computer, or server. However, with the increasing use of virtual systems and cloud provisioning, the notion of physical hardware being tied to a particular system is becoming less common. So an information node is simply an identifiable "system" that the organization runs.

The information node is the lead pattern in a large information group that describes different types of systems. The application is represented by a pattern from the group called **A**PPLICATION **N**ODE. The **S**TAGING **A**REA pattern is also in the same group.

The information node provides an execution environment for the information processes, information services, and information collections. Calls between these components can occur totally within the information node. However, it is also possible for information processes to access information from different information nodes. This capability is provided a specialist pattern within the information service pattern group called **R**EMOTE **I**NFORMATION **S**ERVICE.

Figure 1.8 illustrates this mechanism. The remote information service uses an **I**NFORMATION **R**EQUEST pattern to retrieve information from an information collection located in another information node. The information request pattern consists of two message flows: one from the remote information service to the information node that hosts the information to request the information, and another flowing in the opposite direction to return the requested information.

The information node that receives the request for information routes it to an appropriate information service to extract the information and return a response. In Figure 1.8, this is shown as a **L**OCAL **I**NFORMATION **S**ERVICE—that is, one that uses information collections from the same information node—but it could be another remote information service.

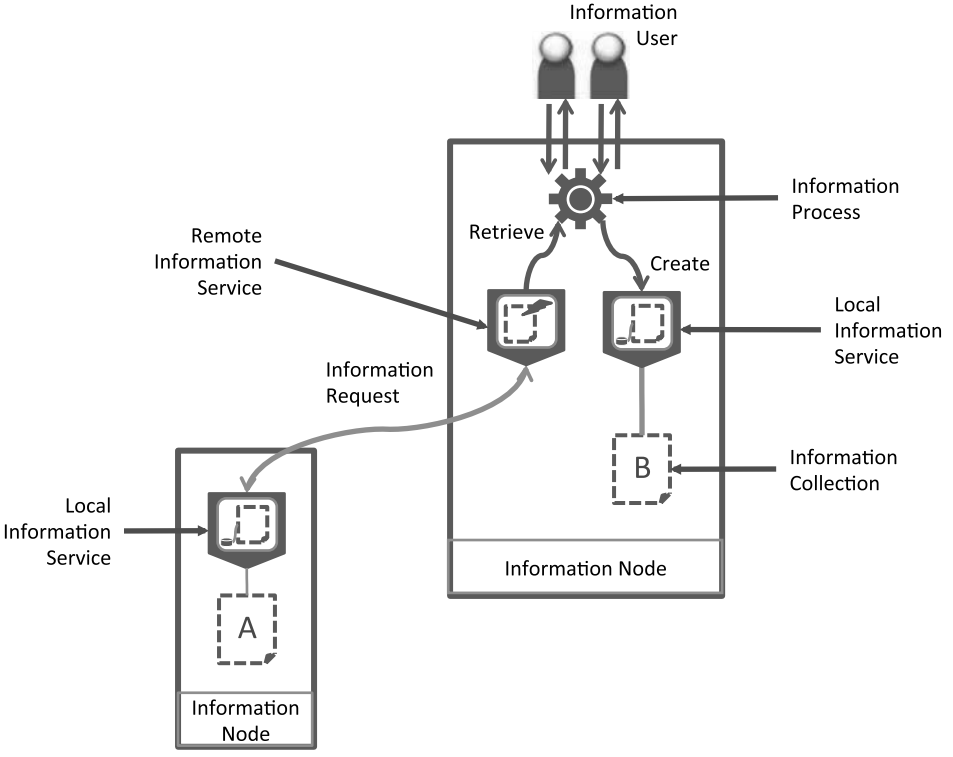


Figure 1.8 Accessing information from a different information node.

The information request pattern retrieves information from its original location on demand. This means both the calling and the called information node must be available at the same time. When information must be copied from one information collection to another—for example, for performance or availability reasons—the information flow pattern is used instead. This introduces another kind of information node called an **I**NFORMATION **B**ROKER that calls remote information services to extract information from one or more information collections, transform it, and store it in other information collections. The effect is that information flows between the information nodes in what we call an information supply chain.

Figure 1.9 illustrates this flow of information. The numbers on the diagram refer to these notes:

1. Here, an information process calls a remote information service to retrieve information from information collection A. Under the covers, the remote information collection uses an information request to contact the information node where information collection A is located.
2. The information process then works with some information users to update the information and store the results in information collection B.
3. Information collection B stores the information in a new entry in the information collection.
4. An information broker now starts an information process to extract the information from information collection B and transform it and save it as a new entry in information collection C.
5. Another information process starts to retrieve the information from information collection C.
6. This process may make changes to the information and update it in information collection C.

This example illustrates how multiple copies of information are created—and also how these copies quickly become slightly different from one another. The differences could be as follows:

- **Superficial—**Such as a reformatting
- **Enriching—**Where additional information is added to the original information
- **Localized—**Where updates made are only relevant to the location where they are made
- **Managed—**Where the best source of information (called the authoritative source) is well known at all times
- **Conflicting—**Where it is hard to know which information collection is the best to use or retrieve the latest information from as changes are coming in to each of the copies in an unpredictable way

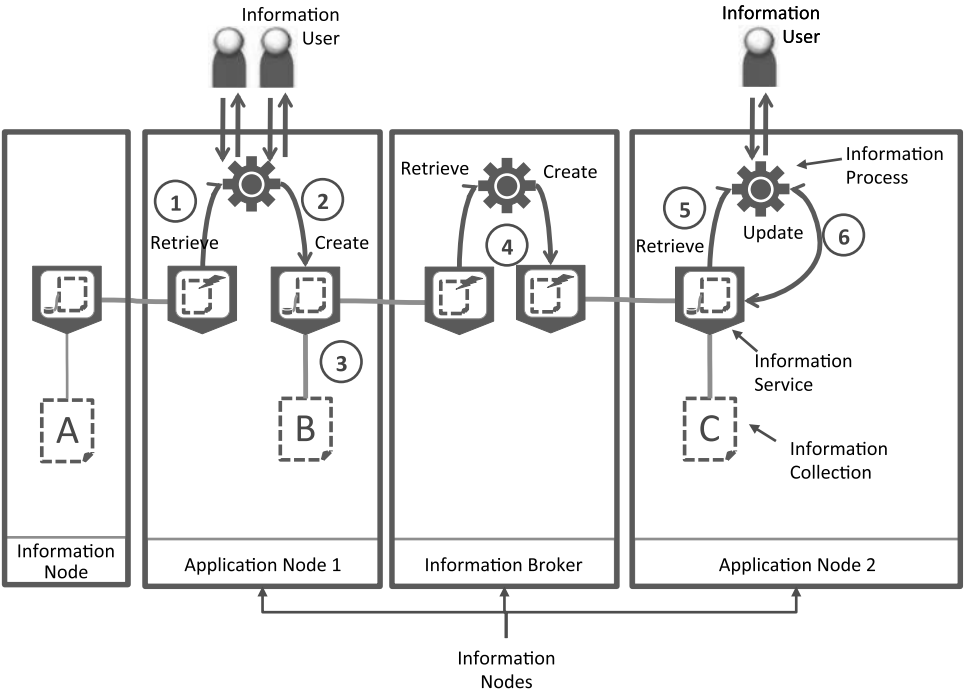


Figure 1.9 Flowing information between systems.

A well-defined information supply chain should avoid having information collections with conflicting differences in them. We aim to minimize the number of copies. Where copies are made, each should have a clear purpose and guidelines on when it should be used. Copies should be synchronized when updates are made, and where differences are unavoidable, there should be at least one copy that is known to have all of the latest information in it.

The **I**NFORMATION **S**UPPLY **C**HAIN pattern is the lead pattern in a pattern group that describes different patterns of information movement between the information collections and how to synchronize the information to avoid conflicting differences. Designing information supply chains is a key challenge for both information architects and solution architects.

Pattern Language Structure

The pattern groups introduced in the previous sections show the breadth of factors an architect needs to consider to achieve a clean, correct, and flexible information design. At the top of the pyramid, you see the organization and the people who work within it. They drive the information strategy. The organization is supported by the information architecture, which is implemented by the information management components. This structure is summarized in Figure 1.10.

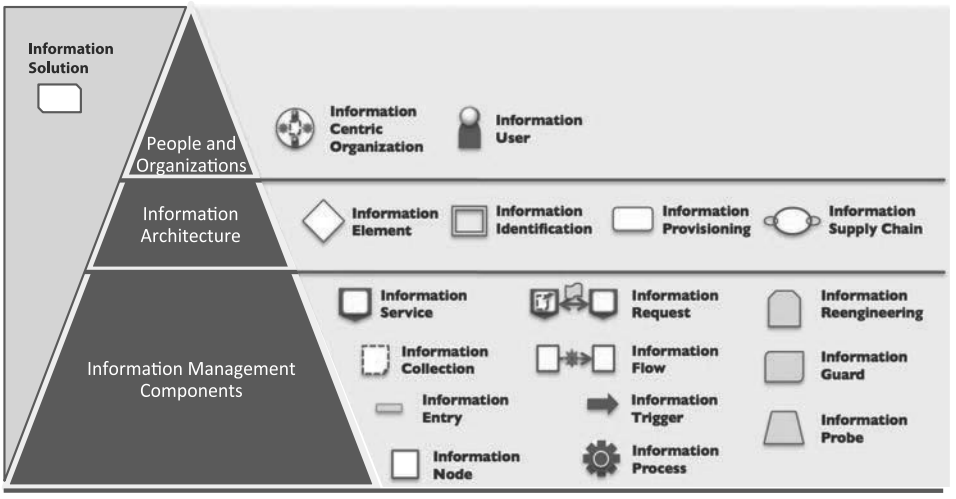
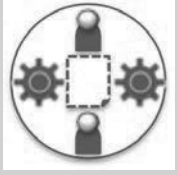


Figure 1.10 Structure of the pattern language for information management.

Chapter 2, "The MCHS Trading Case Study," contains the case study, introducing the patterns through different projects at MCHS Trading. The remaining chapters contain the pattern descriptions and are organized according to the structure in **Figure 1.10**.

Chapter 3, "People and Organizations," has the patterns for the organization and its people (see Table 1.3). These patterns include the information strategy, policy setting, and information governance.

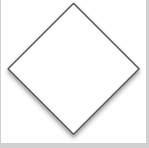
Table 1.3 Pattern Groups in Chapter 3

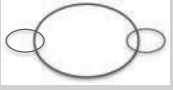
Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **C**ENTRIC **O**RGANIZATION 	An organization needs to make good use of its information to achieve its goals.	Make the management of information a strategic priority. Develop systems and practices that nurture and exploit information to maximum effect.	You are thinking about the holistic approach that your organization should take to information management.

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION USER 	Individuals need access to the organization's information to perform their work.	Classify the people connected with the organization according to their information needs and skills. Then provide user interfaces and reports through which they can access the information as appropriate.	You want to define what types of user roles should be supported by a new information solution.

Information architects develop an understanding of the information needs of an organization and propose best practices for how it should be structured, stored, and managed. Solution architects are responsible for developing IT-based solutions to business problems. These solutions are dependent on information and so the solution architect relies on the information architecture created by the information architect when developing a new solution. Both the information architect and the solution architect use information architecture patterns in their work. These are described in the pattern groups shown in Table 1.4, and are described in more detail in Chapter 4.

Table 1.4 Pattern Groups in Chapter 4

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION ELEMENT 	An organization is looking for the best approach to manage the many kinds of information it has.	Group together related information attributes that follow the same life cycle and manage them appropriately.	You are new to information management and want to familiarize yourself with the types of information an organization has, and how it is managed.
INFORMATION IDENTIFICATION 	An organization does not know what types of information it has, where it is located, how it is managed, and who is responsible for it.	Investigate and document the information requirements and existing support available to the organization.	You want to catalog the information you have and any new requirements. The resulting information is often called metadata.

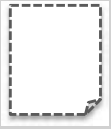
Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **P**ROVISION-ING 	An information process needs information to perform its work.	Information is supplied to the process when it starts, through its user interfaces and through stored information.	You are considering how to provide information to an information process or information user.
INFORMATION **S**UPPLY **C**HAIN 	An organization needs to process information in order to fulfill its purpose. How is the flow of information coordinated throughout the organization's people and systems?	Design and manage well-defined flows of information that start from the points where the information is collected for the organization and links the flows to the places where key consumers receive the information they need.	You are designing how a particular type of information should flow between your systems.

Chapter 5, "Information at Rest," covers the way information is processed within an IT system (see Table 1.5). In this pattern language, a system is called an information node and there is a related pattern group devoted to the various types of systems.

Stored information is accessed through information services that locate the required information and format for the consumer. Information collections are logical groupings of stored information. Often the information is organized consistently within the collection. Approaches for identifying, structuring, locking, and storing information within an information collection are covered in the information entry pattern group.

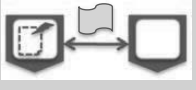
Table 1.5 Pattern Groups in Chapter 5

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **S**ERVICE 	Some information processes need the same information, but may require it to be formatted differently.	Define well-defined interfaces to the information that meet the needs of particular consuming information processes to enable them to create, retrieve, and maintain just the information they need.	You need to decide how an information process will access the information it needs.

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **C**OLLECTION 	Information must be organized so it can be located, accessed, protected, and maintained at a level that is consistent with its value to the organization.	Group related information together into a logical collection and implement information services to access and maintain this information.	You need to classify how the existing stores of information are used or decide how new information should be grouped and stored.
INFORMATION **E**NTRY 	An instance of a type of information needs to be stored in an information collection.	Structure the information collection so that it is made up of a set of information entries. Each information entry stores a single instance of the subject area. Provide capability to manage and iterate over a collection of these archetypal instances.	You are designing how information should be managed within an information collection.
INFORMATION **N**ODE 	What is the appropriate IT infrastructure to host information collections and information processes?	Related information processes and information collections should be hosted together in a server.	You are selecting the type of system to host information and its related processing.

Chapter 6, "Information in Motion," covers the information flow and information request pattern groups for moving information between information nodes (see Table 1.6).

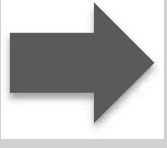
Table 1.6 Pattern Groups in Chapter 6

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **R**EQUEST 	An information process needs to work with information located on a remote information node.	Open a communication link with the remote information node and synchronously exchange the information and associated commands using an agreed protocol.	You want to understand the data that flows between two communicating information nodes.

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **F**LOW 	How do you implement the movement of information between two information nodes?	Use an information trigger to start an information process to control the movement of information. This information process is responsible for extracting the required information from the appropriate sources, reengineering it, and delivering it to the destination information nodes.	You are designing information integration jobs to move information between different systems.

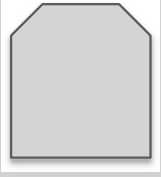
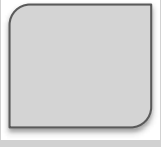
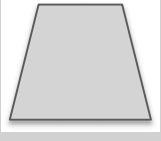
Chapter 7, "Information Processing," covers the different kinds of information processes, along with the information triggers that start them, which are found in a typical organization's IT systems (see Table 1.7).

Table 1.7 Pattern Groups in Chapter 7

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **T**RIGGER 	An information process must be started when a particular event occurs.	When the event is detected, trigger a mechanism that is able to request the initiation of the process on an appropriate information node.	You are considering how to start an information process. This information process may be providing a new business function or moving information between information collections.
INFORMATION **P**ROCESS 	An organization has to process information to support one of its activities.	Formally define and implement the processing for that activity in an information node. Ensure this information node has access to the information it needs.	You want to understand existing processing and or design new processing of information.

Another key concern for organizations with valuable information is how to protect it so it retains its quality and it is not misused or stolen. This is covered in Chapter 8, "Information Protection," (see Table 1.8).

Table 1.8 Pattern Groups in Chapter 8

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **R**EENGINEERING **S**TEP 	An information process is not able to consume the information it needs, as currently exists.	Insert capability to transform the information so it is consumable by the information process.	You need to understand how information can be transformed to meet new requirements.
INFORMATION **G**UARD 	The organization's information needs to be protected from inappropriate use and theft.	Insert mechanisms into the information supply chain to verify that the right people are only using information for authorized purposes.	You need to consider the alternatives for the security and privacy of your information.
INFORMATION **P**ROBE 	The operation of an information supply chain needs to be monitored to ensure it is working properly.	Insert probes into key points in the information supply chain to gather measurements for further analysis.	You need to plan how information management should be monitored.

The protection of information is something that must be designed holistically, considering the welfare of key information at all stages of its lifetime. It is then implemented through the deployment of small components throughout the systems, where each is responsible for protecting an aspect of the information. The patterns of information management break down the aspects of information protection into three pattern groups:

- **INFORMATION REENGINEERING STEP—**These patterns focus on maintaining the quality and format of information.
- **INFORMATION GUARD—**These patterns ensure authorized people and processes are using information for authorized purposes.
- **INFORMATION PROBE—**These patterns are used to monitor the use and movement of information. With these patterns, it is possible to detect issues in the management of information and correct it.

The information protection patterns are used as processing steps in both the information process and information service pattern groups where they transform, protect, or monitor information as it enters the organization; when it is stored; when it is sent between systems; when it is retrieved, updated, and eventually archived and deleted. Individually, they protect a single point in the processing—collectively, they protect the organization's information throughout its entire life cycle.

The final pattern group in Chapter 9, "Solutions for Information Management," covers solutions that tackle different aspects of how information management can be improved. They use the pattern groups described previously as components (see Table 1.9).

Table 1.9 Pattern Groups in Chapter 9

Lead Pattern Name and Icon	Lead Pattern Problem Statement	Lead Pattern Solution Summary	Start Here When...
INFORMATION **S**OLUTION 	An organization recognizes there is a missing capability or a major issue with the way it manages an aspect of its information.	Create a project, or series of projects, to transform the way the information is managed by the organization's people and information systems.	You want to plan changes to your information systems to improve information management.

Summary

This chapter introduced some of the information management challenges an organization faces. Their applications provide the information processes that drive the business. These processes need access to a variety of information to perform their function. This information is distributed and duplicated among the applications and the challenge is to keep this information synchronized while ensuring it is available and suitably structured for all of the organization's needs.

This chapter also introduced the patterns of information management. The patterns of information management are a collection of software design patterns that describe best practices for blending software components together to manage the typical information management challenges that organizations face. These patterns each have a succinct name and icon for use in design discussions. Each pattern also has a tabulated short description called a patlet and a full description that explains when to use it, how it works, and the consequences of using it.

Throughout the pattern language, this book uses a fictitious company called MCHS Trading to illustrate the use of the patterns. The patterns are also grouped together around particular information management topics called pattern groups. Each pattern group has a lead pattern that describes the basic mechanism at work and the rest of the patterns in the group are variations of this basic pattern.

The first lead pattern for a pattern group that was introduced was information provisioning. This explained the layers of components used to provide information to the organization. We then went on to explain how the pattern language is structured and where each of the pattern groups are located in the book.

Now that you have seen the pattern groups in the pattern language, you can choose to navigate directly to the patterns of interest. Alternatively, Chapter 2 describes how MCHS Trading used the patterns to transform its information systems through a series of projects.

This page intentionally left blank

This page intentionally left blank

Index

C

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W-X-Y-Z