



Previews of TDWI course books are provided as an opportunity to see the quality of our material and help you to select the courses that best fit your needs. The previews cannot be printed.

TDWI strives to provide course books that are content-rich and that serve as useful reference documents after a class has ended.

This preview shows selected pages that are representative of the entire course book. The pages shown are not consecutive. The page numbers as they appear in the actual course material are shown at the bottom of each page. All table-of-contents pages are included to illustrate all of the topics covered by a course.



# TDWI Master Data Management Fundamentals

---

The Data Warehousing Institute takes pride in the educational soundness and technical accuracy of all of our courses. Please send us your comments—we'd like to hear from you. Address your feedback to:

email: [info@tdwi.org](mailto:info@tdwi.org)

Publication Date: February 2011

© Copyright 2011 by The Data Warehousing Institute. All rights reserved. No part of this document may be reproduced in any form, or by any means, without written permission from The Data Warehousing Institute.

# TABLE OF CONTENTS

<b>Module 1</b>	<b><i>MDM Concepts .....</i></b>	<b><i>1-1</i></b>
<b>Module 2</b>	<b><i>MDM Processes and Architectures .....</i></b>	<b><i>2-1</i></b>
<b>Module 3</b>	<b><i>Identity Management .....</i></b>	<b><i>3-1</i></b>
<b>Module 4</b>	<b><i>Hierarchy Management .....</i></b>	<b><i>4-1</i></b>
<b>Module 5</b>	<b><i>Implementing and Operating MDM .....</i></b>	<b><i>5-1</i></b>
<b>Module 6</b>	<b><i>Summary and Conclusion .....</i></b>	<b><i>6-1</i></b>
<b>Appendix A</b>	<b><i>Bibliography and References .....</i></b>	<b><i>A-1</i></b>
<b>Appendix B</b>	<b><i>Exercises .....</i></b>	<b><i>B-1</i></b>



---

# Module 1

---

## MDM Concepts

Topic	Page
Defining MDM	1-2
The Need for MDM	1-12
The Basics of MDM	1-16
Dimensions of MDM	1-22
MDM Challenges	1-24

# Defining MDM

## Master Data

Master data refers to the non-transactional data entities of an enterprise – customers, products, parts, services, suppliers, employees, accounts, etc. – that support transactional activities and that are used by many groups and processes throughout the organization.



overlaps with the reference data of an organization  
the context of business transactions and business analysis  
collected & maintained by many different processes  
characterized by redundancy and inconsistency

# Defining MDM

## Master Data

### WHAT IS MASTER DATA?

Master data is a term used to describe the non-transactional data entities of an enterprise – customers, products, parts, services, suppliers, accounts, etc. – that support transactional activities and that are used by many groups and processes throughout the organization.

Master data has these properties:

- It overlaps with reference data of an enterprise.  
*All master data is reference data, but not all reference data is master data.*
- It provides context for business transactions and business analysis.  
*Master data and conformed dimensions have much in common.*
- It is collected and maintained by many different processes.  
*Historically it has been managed in “stovepipe” fashion.*
- It is characterized by redundancy and inconsistency.  
*Thus we have need for master data management.*

### OTHER DEFINITIONS

There is no single, standard definition of master data. In addition to the definition above, these definitions from the experts add some insight into the nature of master data:

“Master data objects are those core business objects used in different applications across the organization, along with their associated metadata, attributes, definitions, roles, connections, and taxonomies.”<sup>1</sup>

“Master data can be defined as the data that has been cleansed, rationalized, and integrated into an enterprise-wide system of record for core business activities.”<sup>2</sup>

Master data is “data describing the people, places, and things involved in an organization’s business ... Master data tend to be grouped into master records, which may include associated reference data.”<sup>3</sup>

“... another term for reference data, which is descriptive data about a business subject area ...”<sup>4</sup>

<sup>1</sup> *Master Data Management*, pp. 5-7, Loshin

<sup>2</sup> *Master Data Management and Customer Data Integration for a Global Enterprise*, pp. 8, Berson & Dubov

<sup>3</sup> *Executing Data Quality Projects*, pp. 299, McGilvray

<sup>4</sup> *Customer Data Integration*, pp. 278, Dyché & Levy

# Defining MDM

## Customer Data Integration (CDI)



# Defining MDM

## Customer Data Integration (CDI)

### CUSTOMER SPECIFIC MDM

Customer Data Integration (CDI) applies MDM techniques and practices to the subject area of customer. CDI encompasses the technology, processes, and services needed to create and maintain accurate, timely, and complete representation of customers across multiple channels, business-lines, and enterprises.

Customer data is a common high-priority focus of MDM because it is especially challenging. Many different business units interact with customers, often collecting and storing customer data redundantly and inconsistently. The impacts can be quite severe – well beyond the simple issue of inefficiency – when redundancy and inconsistency becomes visible to customers. Embarrassing situations are sure to happen when customer communications are poorly coordinated and driven by data of uncertain quality. Beyond embarrassment, there is real risk of regulatory and legal consequences. Consider, for example, the customer who can't successfully opt out of email communications from a company because there is no single record of customer preferences. Each instance for every customer who seeks to opt out is a violation of the CAN-SPAM act.

CDI extends MDM to include practices, technologies, and services specific to the challenges of customer data. Customer data has some unique challenges. In addition to the many data sources, the single common identifier is typically customer name. Two problems arise with using name as an identifier: (1) person names are not assured to be unique, and (2) one individual's name may be expressed in many ways with the variety of formal names, nicknames, initials, etc. Another major challenge with customer data is volatility. One estimate says "2% of records in a customer file become obsolete in one month because customers die, divorce, marry and move."<sup>1</sup>

### CDI DEFINED

In their definitive book *Customer Data Integration: Reaching a Single Version of the Truth*, Jill Dyché and Evan Levy define CDI as "the combination of processes, controls, automation, and skills to standardize and integrate customer data originating from different sources."<sup>2</sup>

Berson and Dubov define CDI as "a comprehensive set of technology components, services, and business processes that create, maintain, and make available an accurate, timely, integrated, and complete view of a customer across lines of business, channels, and business partners."<sup>3</sup>

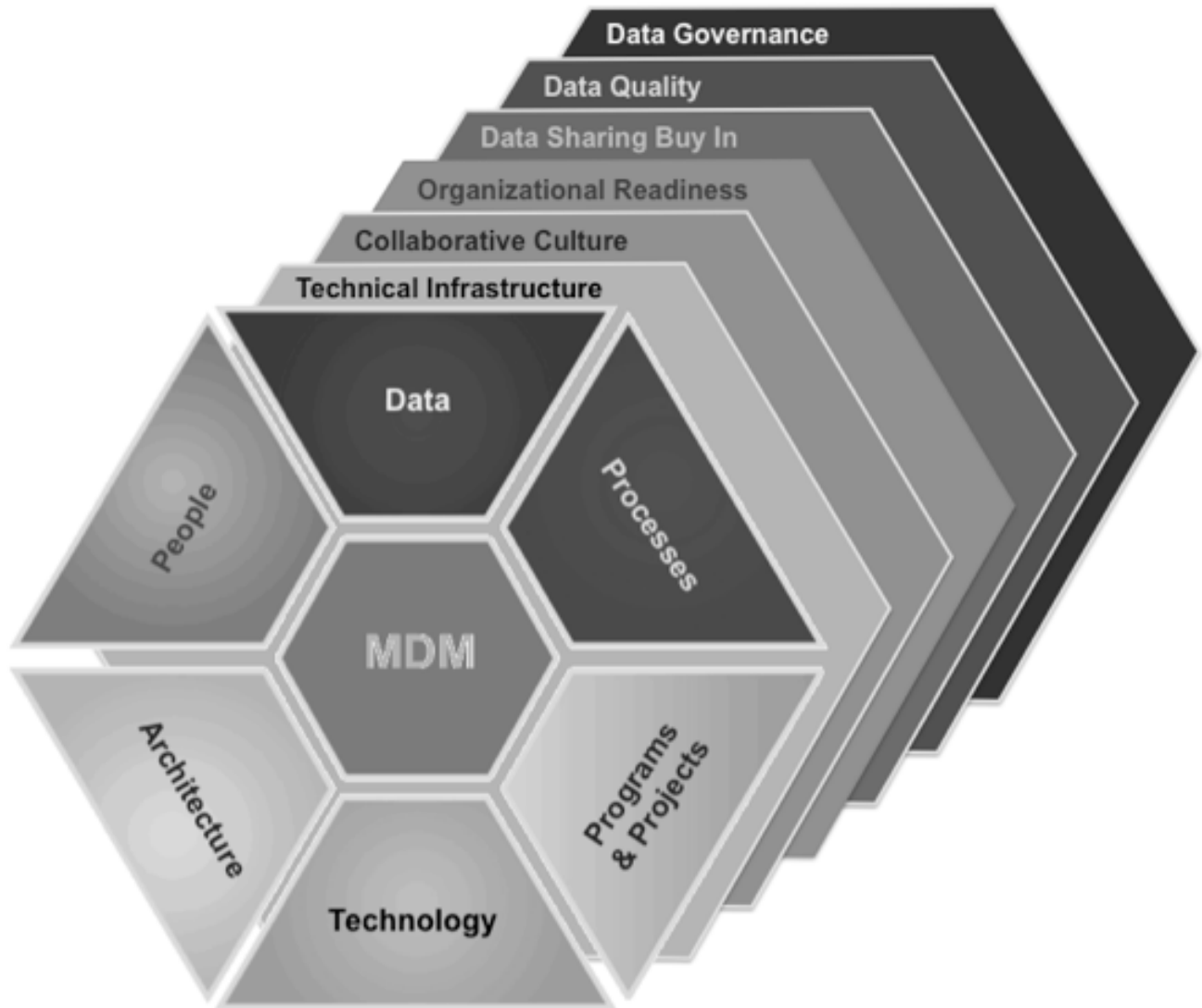
<sup>1</sup> TDWI Report: *Data Quality and the Bottom Line*, pp. 3, Eckerson (<http://download.101com.com/pub/tdwi/Files/DQReport.pdf>)

<sup>2</sup> *Customer Data Integration*, pp. 274, Dyché & Levy

<sup>3</sup> *Master Data Management and Customer Data Integration for a Global Enterprise*, pp. 14, Berson & Dubov

# MDM Challenges

## Prerequisites to Success



---

# MDM Challenges

---

## Prerequisites to Success

### **TECHNICAL INFRASTRUCTURE**

MDM tools are a key part of implementation and operation. Those tools, however, don't stand alone. They need to fit into an existing technical infrastructure that includes both tools and architecture. Services are at the core of MDM, so Service Oriented Architecture (SOA) and service-based technology is important. Data architecture is equally important, as are data management tools for modeling, profiling, metadata management, data quality, data migration, and database management systems. Also consider security and access control as essential infrastructure to implement MDM for privacy-sensitive subjects such as customer and employee data.

### **COLLABORATIVE CULTURE**

MDM is more than just technology and data management. It is substantial change in how information is managed in the business, and it touches nearly every part of the business. Vertically, it affects all levels from executive to administrative. Horizontally, it cuts across many, if not all, business functions. Real collaboration – working together for the benefit of all – is a must for MDM to succeed.

### **ORGANIZATIONAL READINESS**

Much of MDM is people-oriented, making organizational readiness an important success factor. Evaluate readiness in several areas including: clarity of goals, stakeholder awareness and buy-in, data and information architecture, data management policies and practices, issue resolution processes, and change management capabilities.

### **DATA SHARING**

Shared data is a fundamental principle of MDM, which simply can't succeed if technology or human behaviors such as politics, territorialism, and miscommunication inhibit or prevent data sharing.

### **DATA QUALITY**

Integration of bad data is not beneficial to anyone. Perfect data isn't an MDM prerequisite, but attention to data quality including quality measurement, monitoring, and management is essential.

### **DATA GOVERNANCE**

MDM is complex and challenging because it involves data integration, data sharing, collaboration, and quality management. It is complex because it has data, people, process, architecture, technology, and program/project dimensions. The complexity becomes very visible when fitting all of the pieces together and making the right business and data management decisions. The key principles of data governance – decision rights, responsibility, accountability, authority, ownership, stewardship, custodianship – are part of the foundation on which MDM is built.





# Module 2

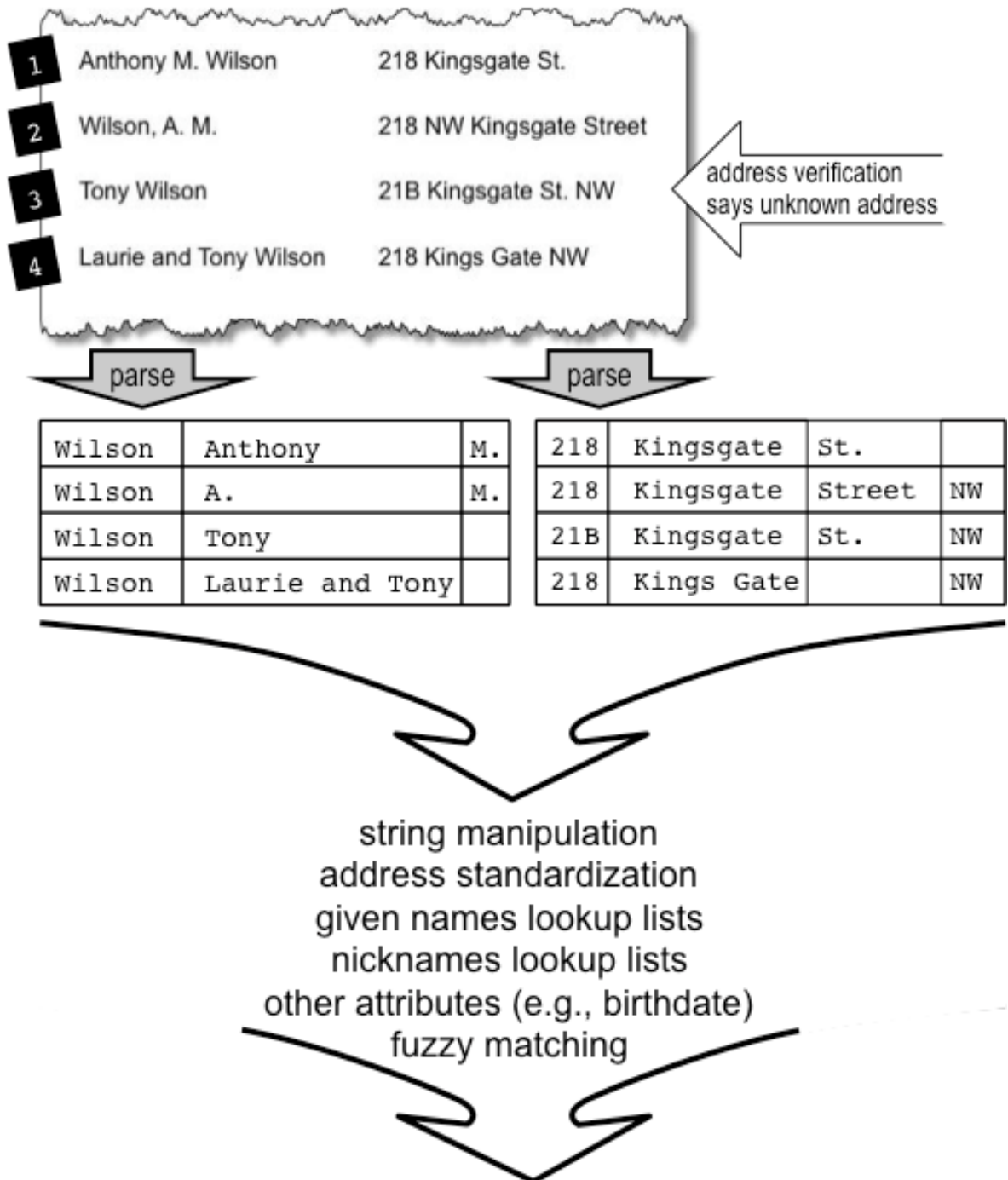
---

## MDM Processes and Architectures

Topic	Page
MDM Processes	2-2
MDM Architectures	2-10

# MDM Processes

## Identity Management



Records 1, 2, and 3 are high probability match. Record 4 contains a match, but it also contains a 2<sup>nd</sup> entity (person) and may require split into two records.

---

# MDM Processes

---

## Identity Management

### **POSITIVE IDENTIFICATION AND UNIQUENESS**

Recall that one of the most basic functions of MDM is to provide an enterprise view of reference entities such as customers, products, employees and accounts. The enterprise view is achieved by recognizing the real-world entity in all of its data instances as a single thing – by identifying it as unique and individual. Identity management activities include matching and identity resolution.

Identity matching involves recognition of individuals (individual customers, suppliers, accounts, employees, etc.) to support positive identification. Recognition of common identity often uses complex logic involving several data elements and algorithms for semantic similarities and match probability.

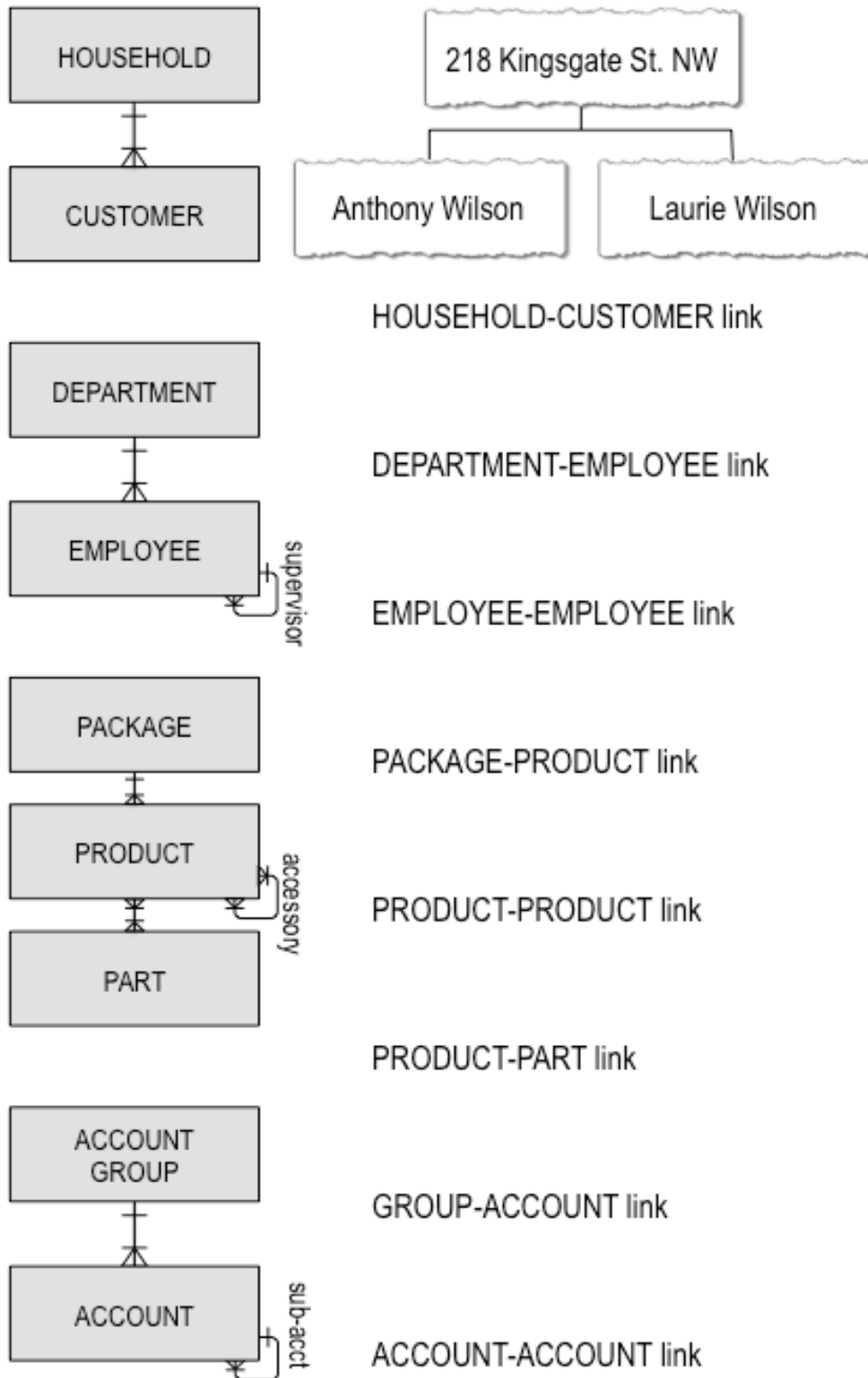
Identity resolution determines what actions to take when multiple records are matched and determined to represent a single individual. Common actions include:

- data linkage, which retains multiple records and registers the association among them
- data consolidation, which combines data from multiple sources to create a single master record.

Identity Management is described in greater detail in *Module Three; Identity Management*.

# MDM Processes

## Hierarchy Management



---

# MDM Processes

---

## Hierarchy Management

### **RELATIONSHIPS OF MASTER DATA RECORDS**

Remember that the second basic function of MDM is to provide an enterprise view of the relationships among reference entities. This function is known as hierarchy management - the ability to define and store relationships between records in the master data resource. The goal is to find related records and link them to constitute the best enterprise-perspective version of data relationships that is available.

There is very strong correspondence between enterprise hierarchy management and the concept of conformed dimensions as it is applied in data warehousing and business intelligence. Conformed dimensions, however, support reporting and analysis only. Enterprise hierarchy management provides services to update and maintain hierarchies as well as for access and reporting – ideally a single point of maintenance that eliminates the need to maintain and synchronize redundant versions of the relationships.

Hierarchy management involves identification and matching, entity and record linking and consolidation, and versioning. Hierarchy management is described in greater detail in *Module Four; Hierarchy Management*.





# Module 3

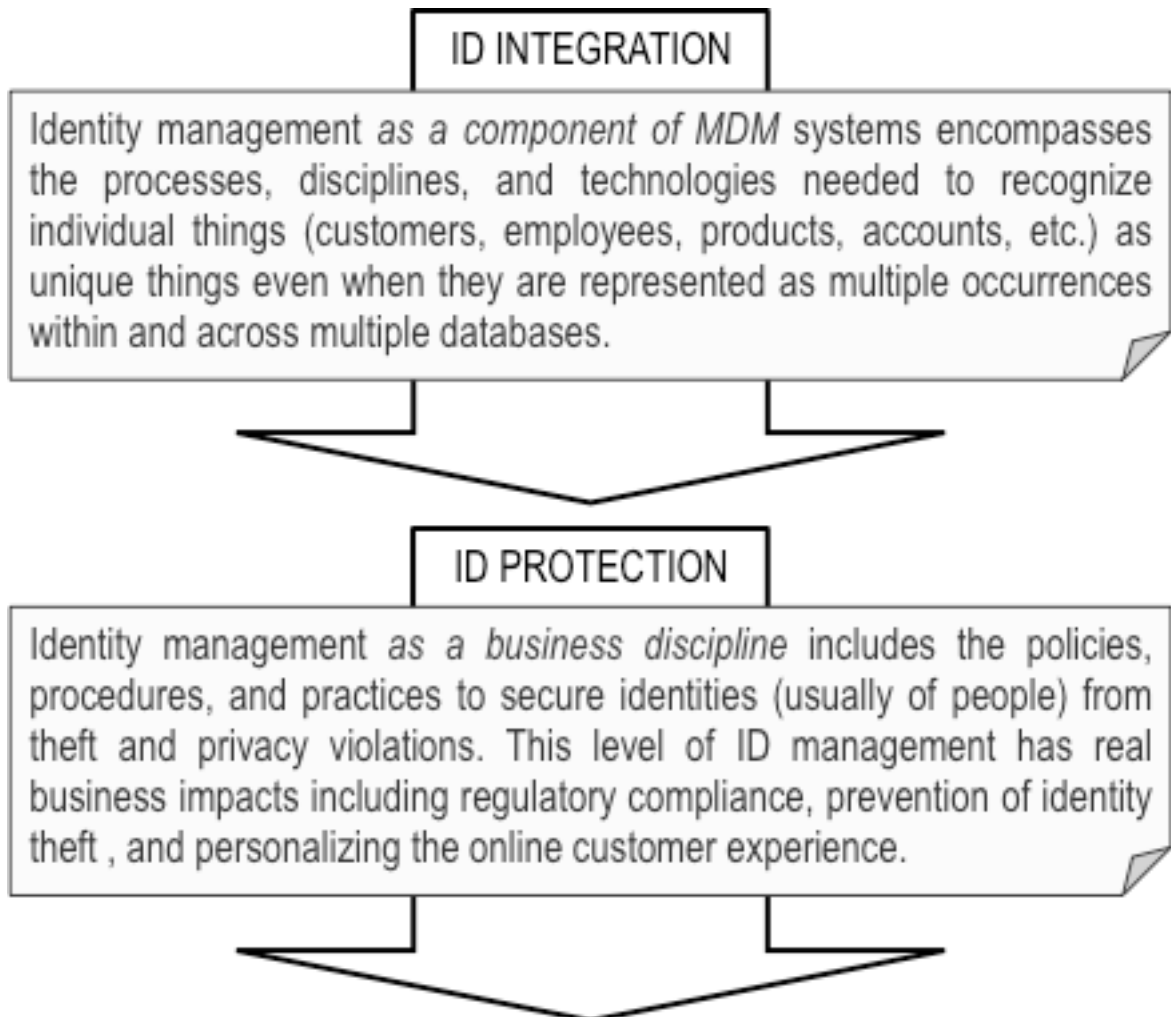
---

## Identity Management

Topic	Page
Identity Management Defined	3-2
Identity Management Functions	3-4
Identity Management in MDM Architecture	3-12

# Identity Management Defined

## Uniqueness of Entities



---

# Identity Management Defined

---

## Uniqueness of Entities

### **IDENTITY INTEGRATION**

Identity management as a component of MDM systems encompasses the processes, disciplines, and technologies needed to recognize individual things (customers, employees, products, accounts, etc.) as unique things even when they are represented as multiple occurrences within and across multiple databases. The purpose is to integrate disparate identities by first matching, then linking, merging, or consolidating multiple records.

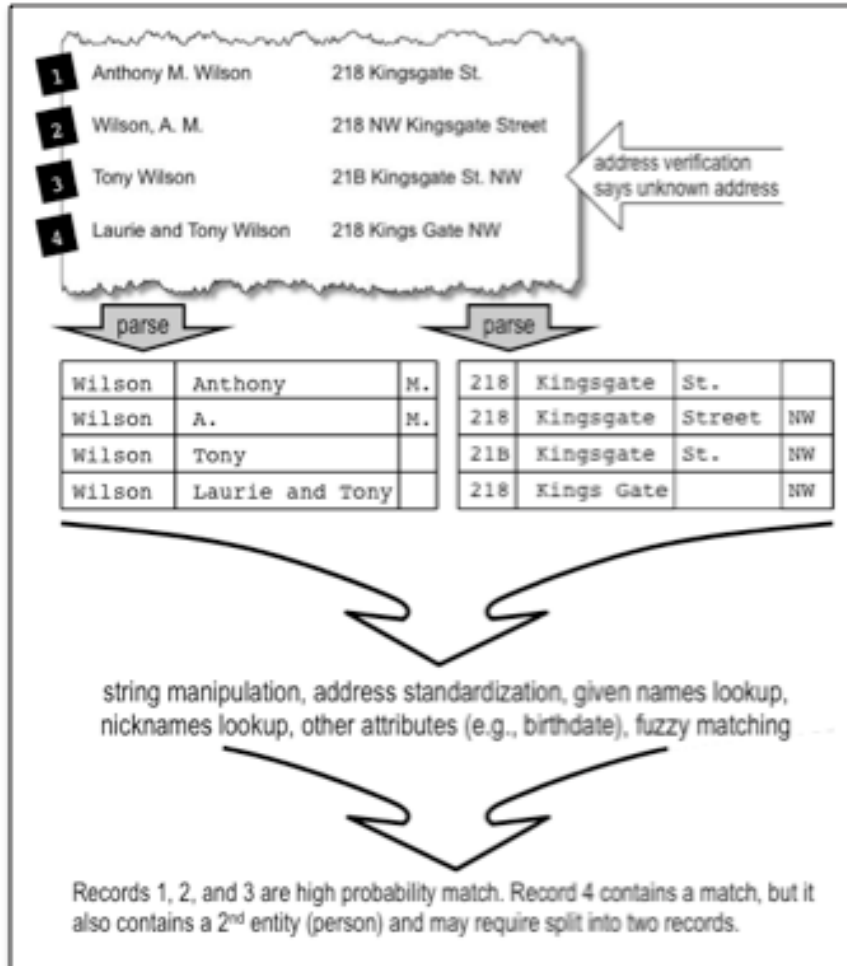
### **IDENTITY PROTECTION**

Identity management as a business discipline builds upon identity integration and includes the policies, procedures, and practices to secure identities (usually of people) from theft and privacy violations. Berson and Dubov define identity management in this context as “an organizing principle, a framework, and a set of technologies designed to manage the flow, consumption, security, integrity, and privacy of identity ...”<sup>1</sup> This level of identity management has real business impacts that include compliance with privacy regulations, prevention of identity theft , and personalizing the online customer experience.

<sup>1</sup> *Master Data Management and Customer Data Integration for a Global Enterprise*, pp. 14, Berson & Dubov

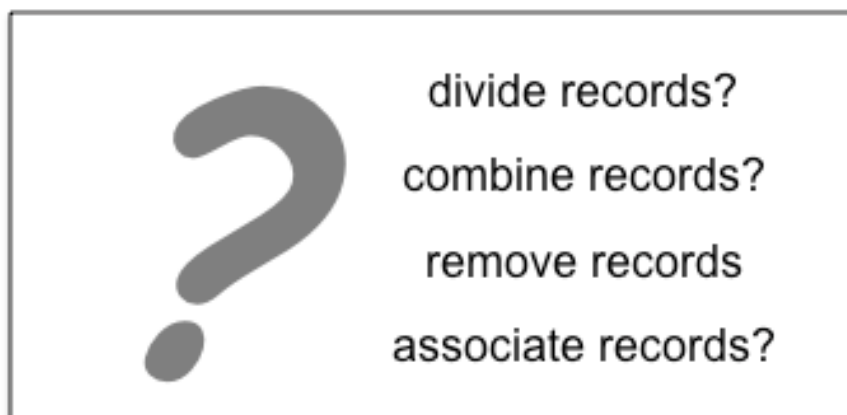
# Identity Management Functions

## Search and Resolution



### SEARCH

- find similar records
- matching & probability
- one-at-a-time (real time)
- batch (cleanse & load)



### RESOLVE

- split overloaded records
- merge and consolidate
- de-duplicate
- link records

---

# Identity Management Functions

---

## Search and Resolution

### **ENTITY REDUNDANCY AND OVERLAP**

The search function in identity management is the work of searching data to find records that appear to be different representations of the same entity. Search is performed using the matching activity described earlier in the course. Record searching typically occurs in three different situations of managing master data:

- In the design and development process as a means to understand the data and establish rules and probability algorithms needed for automated search and resolution. “The first stage is one of discovery and combines data profiling with a manual review of the data.”<sup>1</sup> The discovery activity is needed to prepare for routine search and resolve activities on a continuing basis.
- In MDM implementation processes where batch search and record matching is needed to cleanse and prepare data for loading into a master data hub.
- In day-to-day operations when a single record is submitted for insert or update. On a one-at-a-time basis, MDM matching services assure integrity of identity, establish correct record links, and prevent accidental duplication of records.

<sup>1</sup> *Master Data Management*, pp. 52, Loshin





# Module 4

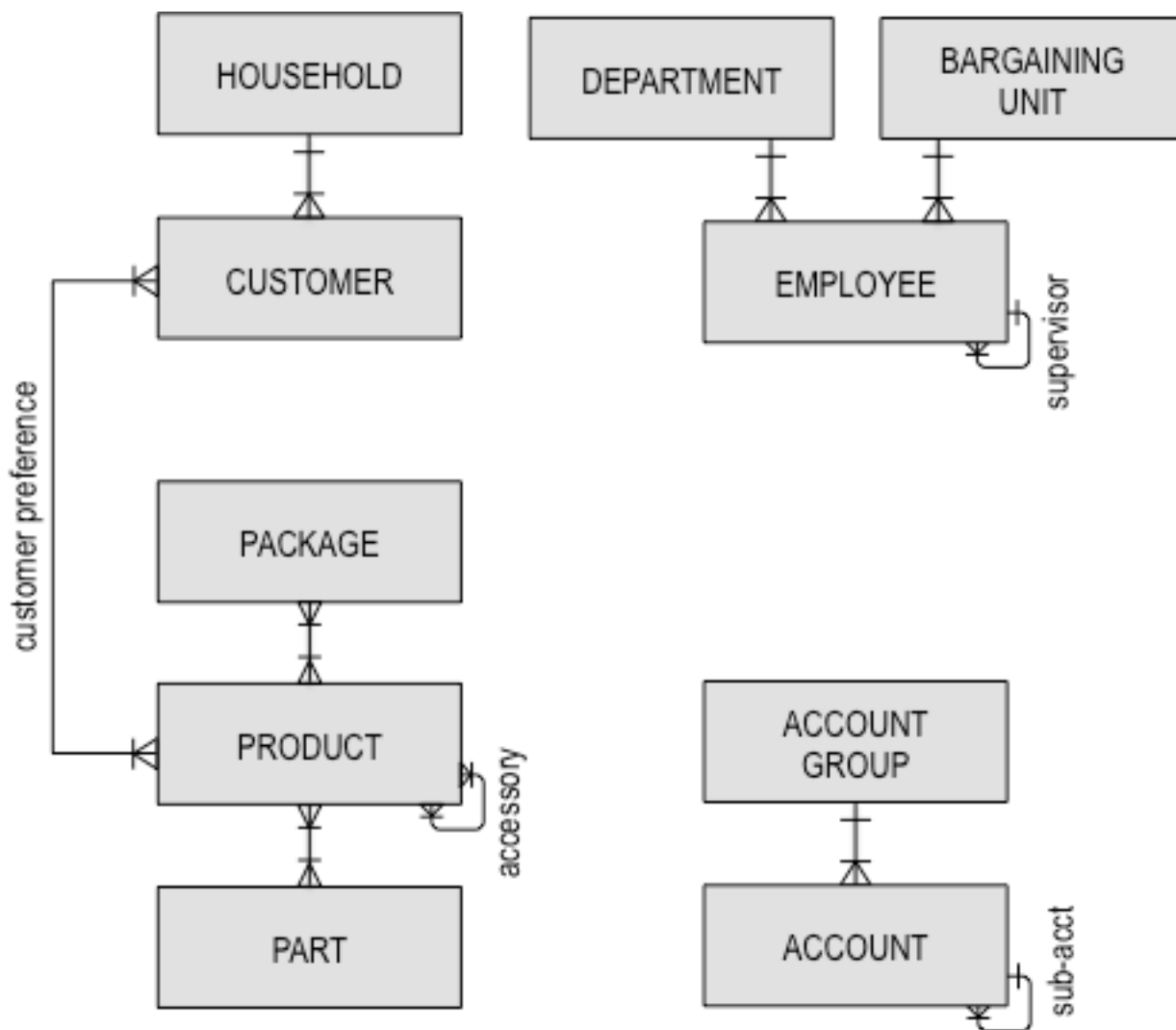
---

## Hierarchy Management

<b>Topic</b>	<b>Page</b>
Hierarchy Management Defined	4-2
Hierarchy Management Functions	4-6
Hierarchy Management in MDM Architecture	4-14

# Hierarchy Management Defined

## Hierarchies as Master Data Objects



# Hierarchy Management Defined

## Hierarchies as Master Data Objects

### MASTER DATA OBJECTS

Many of us view a data model and see entities as “things” in the model and relationships only as connections between the things. In hierarchy management it is important to consider relationships as things that are equally as significant as entities. The data model on the facing page shows ten entities, but it contains twenty master data objects. Listed alphabetically by type they are:

#### Master Reference Objects

Account  
Account Group  
Bargaining Unit  
Customer  
Department  
Employee  
Household  
Package  
Part  
Product

#### Master Hierarchy Objects

Account Group Account  
Customer Household  
Customer Preference  
Employee Bargaining Unit  
Employee Department  
Employee Supervisor  
Product Accessory  
Product Package  
Product Part  
Subaccount

Note that the relationships are named with nouns – not the typical naming convention for relationships in an entity-relationship data model. The names are intended to describe objects – things that are managed as master data. Direction in naming is somewhat arbitrary – *Customer Household* instead of *Household Customer* – with best effort to name in the way that makes the most business sense.

### MULTIPLE HIERARCHIES

Multiple hierarchies are common and easily implemented. Here we see three distinct organizational hierarchies for Employee – one based on department or work assignment, another for collective bargaining, and a third based on supervisory relationships.

### MULTI-SUBJECT RELATIONSHIPS

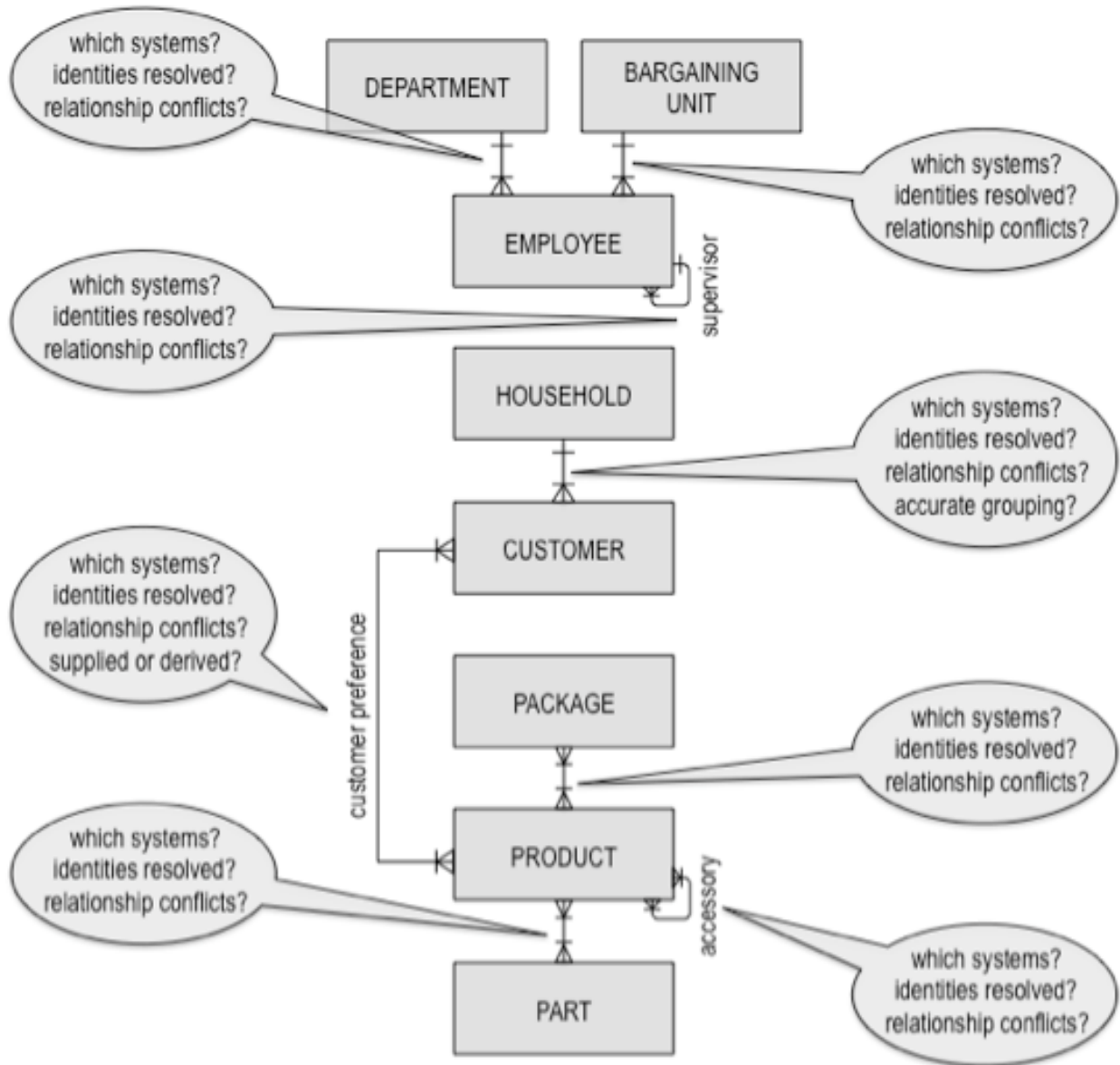
Relationships across subject areas are less common in MDM but are quite acceptable if they provide information of interest across many systems or business functions. Customer Preference is an example.

### MANY-TO-MANY RELATIONSHIPS

By strict definition many-to-many relationships are not true hierarchy. But they do occur in business and actually represent two hierarchies. The Product Part relationship describes (1) which parts are available for a product, and (2) which products use a part. Both can be managed as a single master data object.

# Hierarchy Management Functions

## Identification and Matching



# Hierarchy Management Functions

---

## Identification and Matching

### **RELATIONSHIP REDUNDANCY AND OVERLAP**

Identification and matching in hierarchy management is similar to the search function in identity management. Where identity management seeks records that appear to be different representations of the same entity, hierarchy management seeks multiple representations of the same relationship. While the goal is similar, the matching activities differ from those of identity matching in several ways:

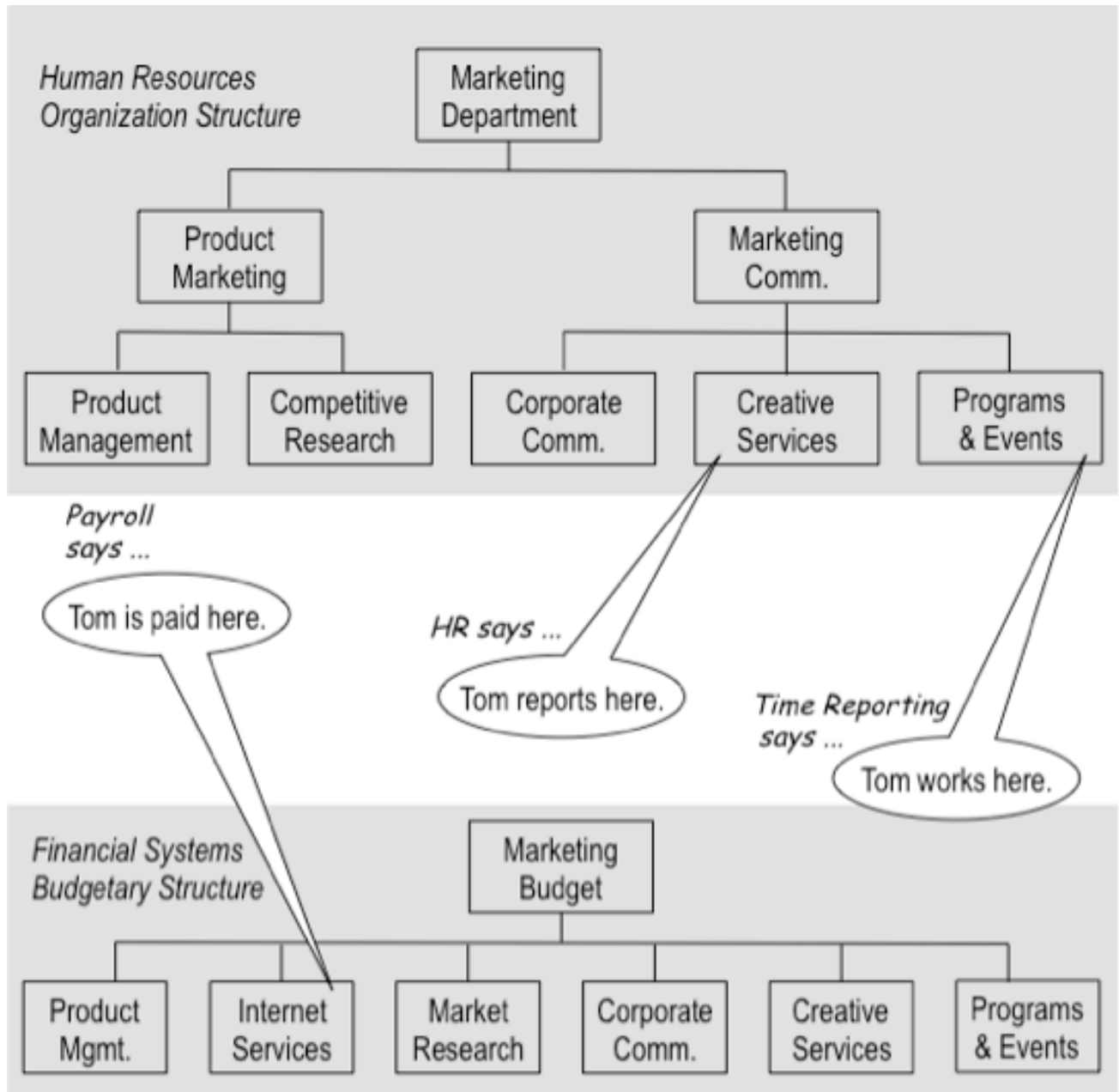
- Hierarchy matching is more precise and less “fuzzy” in hierarchy matching than entity matching. The focus of matching is primary and foreign keys, which typically don’t have the idiosyncrasies of data elements such as name and address.
- Hierarchy matching is difficult to achieve until identity resolution is complete. Prerequisites for hierarchy matching include:
  - ▣ Known source systems and databases for master reference objects – the entities in MDM.
  - ▣ Identities fully resolved for master reference objects.
  - ▣ Key mapping from master data keys to source keys.
  - ▣ Confident and consistent grouping when MDM processes create aggregate objects.
  - ▣ Known source of data for multi-subject relationships. Are they supplied by operational systems or derived by MDM processing?
- Hierarchies are more volatile than identities. Matching is a continuous process and re-matching must take place whenever a foreign key relationship change occurs in any source system or database.

### **RELATIONSHIP CONFLICT**

Relationship conflict exists when various source systems or databases describe the relationships of entities differently. Matching activities must recognize conflicts, which are resolved by consolidation processing.

# Hierarchy Management Functions

## Consolidation



# Hierarchy Management Functions

## Consolidation

### CONFLICTING HIERARCHIES

Hierarchy conflicts occur when various information systems and business processes have different views of the relationships among things. Each view may be correct from a business process perspective. The challenge is to resolve the differences to create enterprise perspective. For entities the enterprise view is always a single authoritative record of an entity. A single authoritative view of relationships may not be the most practical solution.

### RESOLVING CONFLICTS

Common forms of hierarchy conflict include variation in:

- Depth of hierarchy – In the example, HR has three levels and budgeting uses only two.
- Structure of hierarchy – In the example, the bottom level is nearly identical but budgeting contains one additional item.
- Placement in the hierarchy – In the example, an employee “Tom” is viewed differently by each of three business processes.

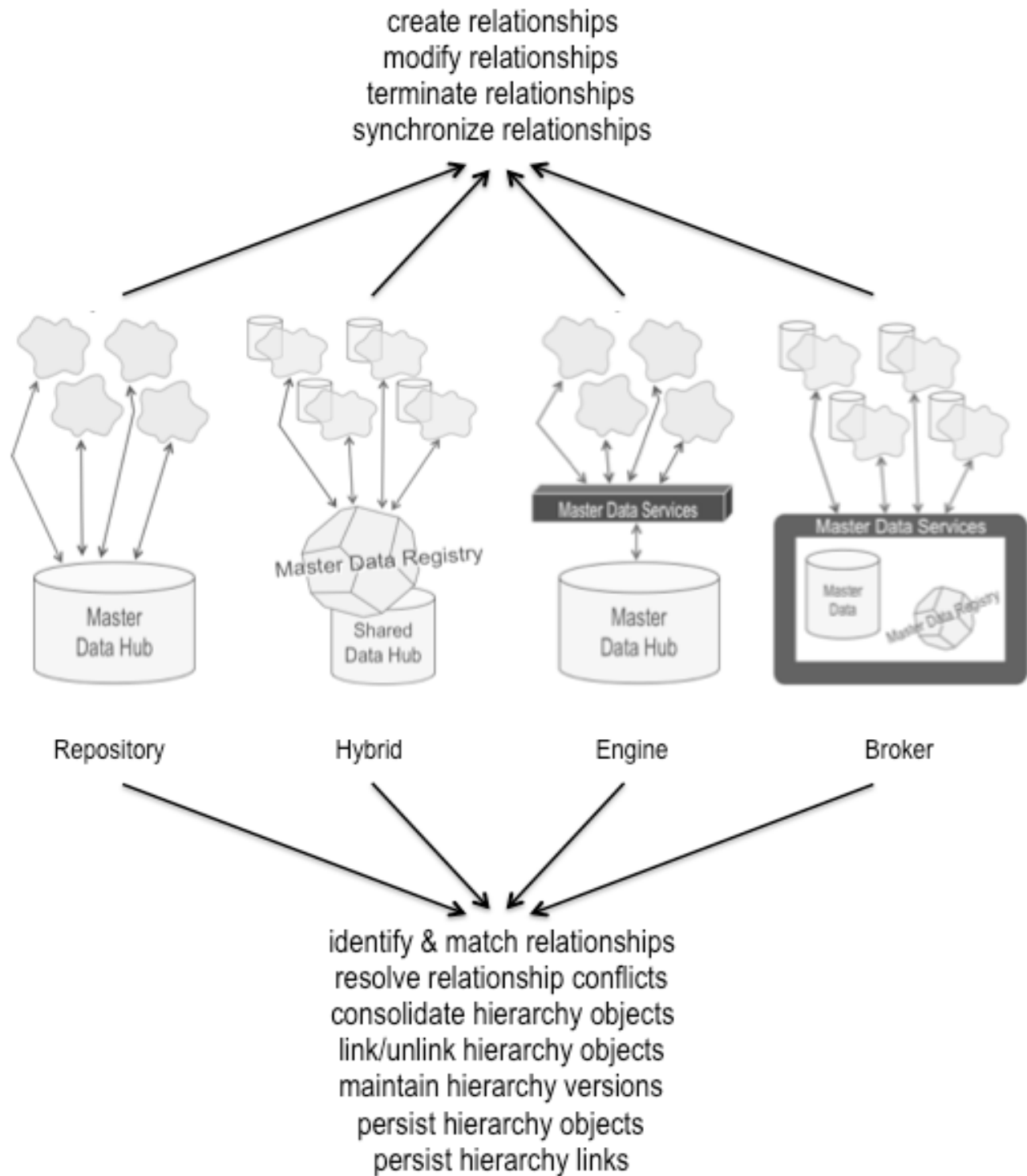
Conflicts may be resolved and relationships consolidated in several ways:

Kind of Conflict	Resolution Method	In the Example
Depth of Hierarchy	Use structure with most levels	Use the HR structure
	Use the structure with fewest levels	Use the budgetary structure
Different Structure	Choose the most widely used	HR or budget?
	Implement two hierarchies	HR and budget
	Blend the structures	Add internet services to HR
Different Placement	Choose the most widely used	Payroll, HR, or time reporting?
	Choose the most stable relationship	Payroll, HR, or time reporting?
	Implement multiple hierarchies	Payroll, HR, or time reporting?

When choosing methods for hierarchy consolidation, consider the implications of multiple hierarchies vs. single hierarchy, highly granular vs. aggregated, etc. Each involves trade-offs of information value, ease of use, complexity, and cost.

# Hierarchy Management in MDM Architecture

## Hierarchies and Hub



---

# Hierarchy Management in MDM Architecture

---

## Hierarchies and Hub

### **ARCHITECTURE REQUIREMENTS**

Hierarchy management needs to have a data store for persistence of hierarchy objects and hierarchy links, thus registry-only architecture is not a good fit. Persistent hierarchy objects are stored in a master data hub, with a component in each of the remaining four approaches: repository, registry-repository hybrid, MDM engine, and MDM broker.

### **APPLICATION CAPABILITIES**

With any of the four architectures, hierarchy management handles the enterprise view of relationships without overriding application views. Applications are able to create, modify, and terminate relationships locally. They may also use MDM services to synchronize locally stored relationships with the enterprise view.

### **APPLICATION RESPONSIBILITIES**

To support hierarchy management, applications are responsible to make their locally maintained relationships known to the MDM system. This may involve either push or pull methods of data exchange. If the MDM system pulls data, then the application needs only to allow MDM access to its internal data. If the application pushes data to MDM, then relatively small application changes are needed.

### **MDM FUNCTIONS**

MDM (with any of four architectures) provides functions or services to:

- Identify relationships and match redundant or overlapping relationships from multiple sources
- Resolve conflicting relationships among multiple sources
- Consolidate relationship data from multiple sources to create singular and authoritative hierarchy objects
- Link and unlink hierarchy objects to maintain their associations
- Maintain multiple versions of hierarchies as required
- Persist hierarchy objects and links in the master data hub for enterprise-wide use





# Module 5

## Implementing and Operating MDM

<b>Topic</b>	<b>Page</b>
Roles and Responsibilities	5-2
Data Governance for MDM	5-4
Data Quality and MDM	5-10
Data Modeling and MDM	5-20
The MDM Program	5-22
MDM Projects	5-26
MDM Operations	5-28

---

# Roles and Responsibilities

---

## Stakeholders



# Roles and Responsibilities

## Stakeholders

### PEOPLE AND MDM

The work of MDM depends on people at least as much as on technology. The technology automates and performs repetitive tasks of data matching, linking, consolidation, synchronization, etc. But those are only the mechanical aspects of MDM. The organizational and business aspects of MDM are strongly connected with stakeholder roles and responsibilities. Key stakeholder groups include those who participate in:

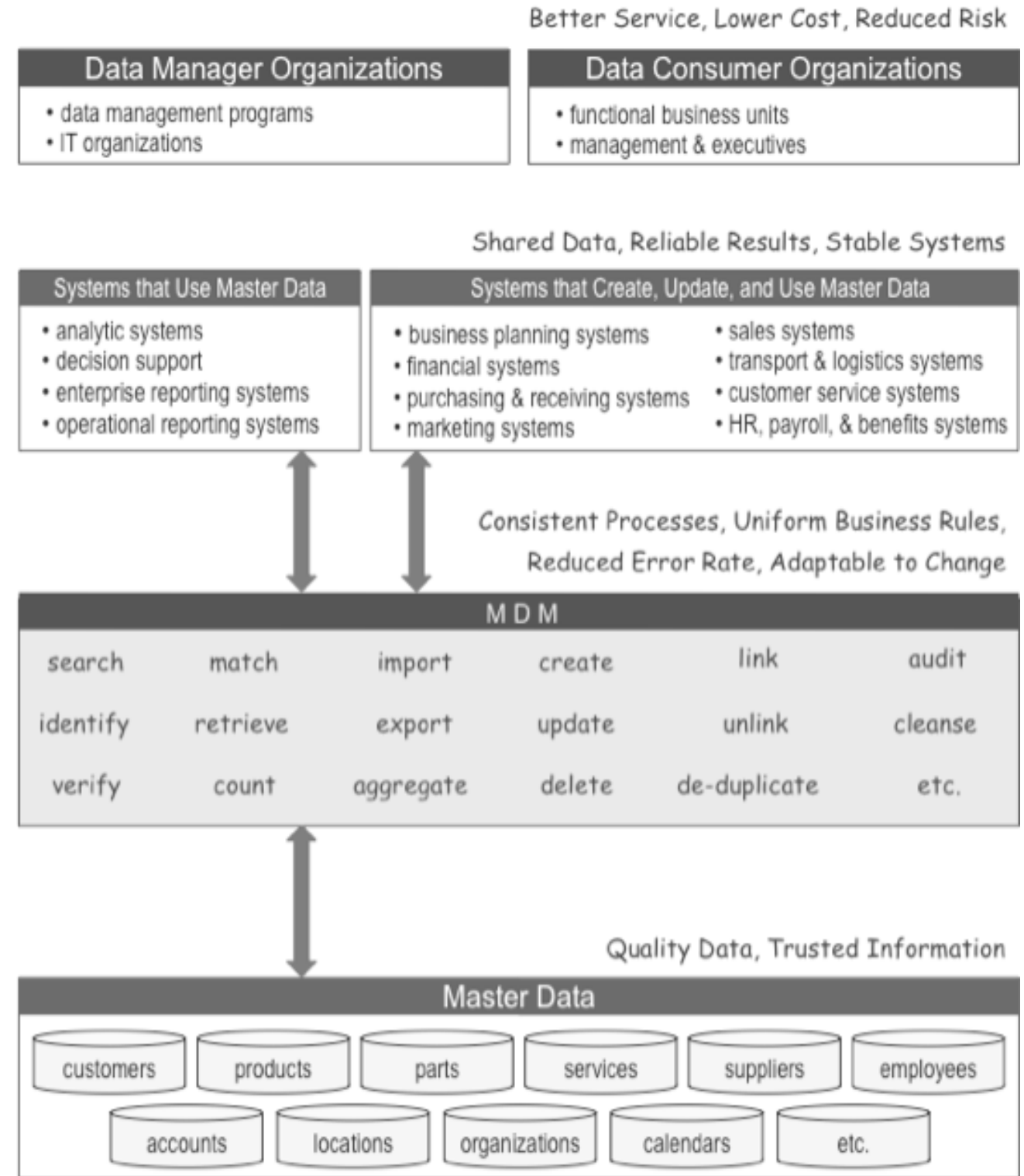
- Program Management
- MDM Development
- MDM Operations
- Data Governance
- Provisioning of Data
- Consuming of Data

The table below summarizes some of the key roles and responsibilities of the stakeholders at three levels – management, implementation, and operation of MDM.

<i>Group</i>	<i>Role</i>	Management	Implementation	Operation
Program Management	Sponsor	funding, support, political will	issue resolution	issue resolution, sustainability
	Program Mgr.	vision, strategy, planning	coordination, issue resolution	quality of service
Development & Operations	Architects	components, roles, standards	conceptual models	
	Analysts		business rules, system needs	
	Developers		design, logical modeling, implement & test components	support?
	DBAs		physical modeling, database implementation	capacity, security, backup, performance
	QA/QC	test criteria	test cases, test tracking	measure & monitor quality
	Operators		testing and acceptance	MDM services availability
Data Governance	Owners	support	knowledge & decisions	authority & accountability
	Stewards	facilitation	consensus, metadata	facilitation, metadata
	Custodians	subject expertise	availability, security, reliability	availability, security, reliability
Information Workers	Providers		Subject expertise	data creation, quality focus
	Consumers		Subject expertise	Data use and feedback

# The MDM Program

## The Business Case



---

# The MDM Program

---

## The Business Case

### **MULTI-LEVEL MDM BENEFITS**

The benefits of MDM occur at several levels

- at the master data asset
- at master data services
- at the systems where master data is created and used
- at data-dependent business functions

While real benefits exist at each of these levels, it is difficult to describe a traditional business case of meeting business requirements. According to Dyché and Levy, business requirements “are a bit overrated when it comes to...MDM deployments” and that “the requirements of individual business users cede to the processing requirements of the applications that need access”<sup>1</sup> to master data. The business case is best described by illustrating the chain from information quality to direct business value.

### **DIRECT BUSINESS BENEFITS**

Business benefits at data-dependent business functions include:

- Quality of service improvements
- Cost savings
- Risk reduction

### **INFORMATION SYSTEMS BENEFITS**

System benefits related to master data include:

- Data sharing
- Consistency and reliability
- Greater stability of systems

### **DATA GOVERNANCE BENEFITS**

Master data services provide governance benefits of:

- Consistent processes
- Uniform business rules
- Reduced error rate
- Adaptability

### **INFORMATION QUALITY BENEFITS**

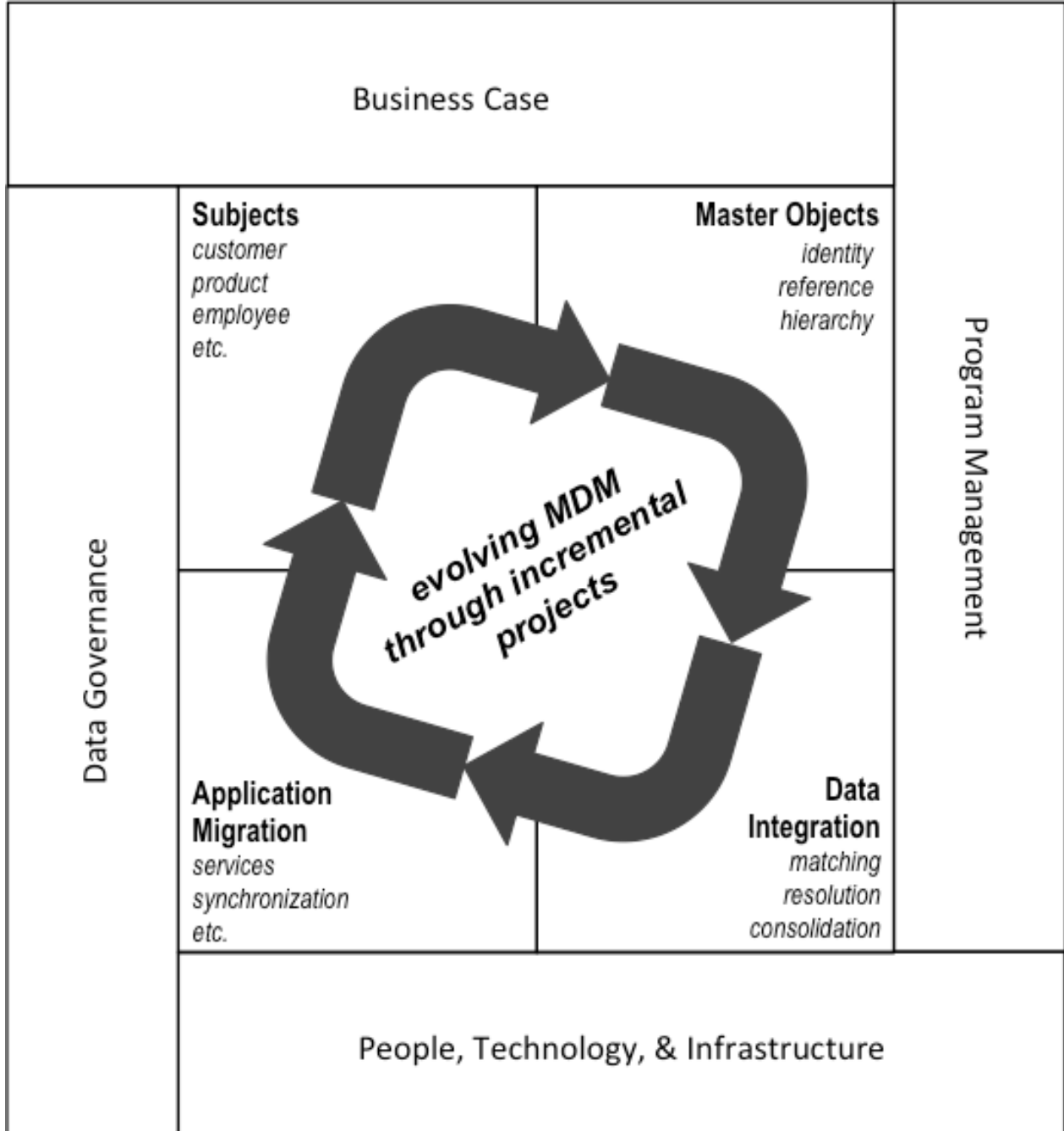
The master data asset provides the foundation benefits of:

- Managed data quality
- Trustworthy information

<sup>1</sup> *Ten Mistakes to Avoid When Planning Your CDI/MDM Project*, Dyché and Levy  
(<http://tdwi.org/research/2006/08/ten-mistakes-to-avoid-when-planning-your-cdi-mdm-project.aspx?tc=page0>)

# MDM Projects

## Implementation in a Program Framework



# MDM Projects

## Implementation in a Program Framework

### FRAMEWORK

Ultimately, MDM is implemented via projects. It is important to note that “projects” is plural. You don’t get to MDM with a single project. Many (and sometimes simultaneous) projects are necessary. The integration goal of MDM demands that project results – data, services, processes, and uses – must all fit together regardless of the numbers and timing of the projects. Several of the topics that we’ve already discussed combine to form a framework within which project cohesion is possible:

- The business case establishes common goals for multiple projects.
- Program management provides direction and coordination.
- Data governance brings enterprise-level data policy and decisions.
- Shared people, technology, and infrastructure form a solid foundation for multiple, interdependent projects.

### KINDS OF PROJECTS

Within the framework you may undertake many different projects to build, evolve, and grow MDM. Early projects, of course, are needed to define the architecture and to implement technology infrastructure. With architecture and technology in place, you’ll find recurring projects to:

- assess master data in its pre-MDM state
- integrate data from source systems into the MDM environment
- migrate applications from local data management to use of master data services

It is almost certain that a separate set of projects is needed for each data subject – customer, product, etc. And it is likely that you’ll need multiple projects for a single subject; you may not want to tackle all customer data sources at once. You may also find it practical to separate reference object consolidation from hierarchy management – to resolve identity and data elements before addressing relationships.

### PROJECT GUIDELINES

These guidelines may help to plan the right collection and sequence of projects for your MDM initiative:

- Start with a single subject. Customer is a common first choice.
- Don’t compromise on data quality even if you must defer problem data sources.
- Reliable identity matching comes ahead of consolidation and synchronization.
- Entity consolidation preceded hierarchy management.





# Module 6

## Summary and Conclusion

Topic	Page
Summary of Key Points	6-2
References and Resources	6-4





# Appendix A

---

## Bibliography and References



# References and Resources

---

## To Learn More

*Customer Data Integration: Reaching a Single Version of the Truth*, Dyché and Levy  
John Wiley & Sons, 2006

*Data Quality Assessment*, Maydanchik  
Technics Publications, 2007

*Data Sharing: Using a Common Data Architecture*, Brackett  
John Wiley & Sons, 1994

*Data Strategy*, Adelman, Moss & Abai  
Addison-Wesley, 2005

*Executing Data Quality Projects*, McGilvray  
Morgan Kaufman, 2008

*Information-Driven Business: How to Manage Data and Information for Maximum Advantage*,  
Hillard  
John Wiley & Sons, 2010

*Managing Your Business Data*, Kushner & Villar  
Racom Books, 2009

*Master Data Management*, Loshin  
Morgan Kaufman, 2009

*Master Data Management and Customer Data Integration for a Global Enterprise*,  
Berson and Dubov  
McGraw Hill, 2007

*The Practitioner's Guide to Data Quality Improvement*, Loshin  
Elsevier, 2011

*Three Dimensional Analysis: Data Profiling Techniques*, Lindsey  
Data Profiling LLC, 2008





---

# Appendix B

## Exercises

---

<b>Exercise</b>	<b>Page</b>
Exercise 1: MDM Readiness	B-2
Exercise 2: MDM Architecture	B-4
Exercise 3: Identity and Consolidation	B-6
Exercise 4: MDM Roles and Skills	B-8
Exercise 5: Readiness Revisited	B-10

# Exercise 1: MDM Readiness

---

## Exercise Instructions

### **READINESS FACTORS**

Use the worksheet on the facing page to evaluate your organization's readiness for MDM in six categories. The ratings are highly subjective and based on your beliefs, opinions, and experiences. Do not consider this exercise as a formal readiness assessment. It is simply a tool to facilitate discussion.

When you've completed the ratings calculate an overall score using the numeric values in the column headings to total all of your responses. (Every item has a value between 1 and 5. Maximum possible score is 35, minimum is 5.)

When the assessment worksheet is completed we'll discuss risk areas and what can be done to mitigate the risks. We'll also discuss strengths and how they can be leveraged for MDM success.

# Exercise 1: MDM Readiness

## Worksheet

Respond to each statement below based on your individual beliefs.

	1	2	3	4	5
Data Governance	none	getting started	some data stewards	pretty good governance	effective data governance
Data Quality Practices	none	informal	pockets of good practices	wide spread good practices	enterprise managed
Source Data Quality	really bad	poor	adequate	good	excellent
Data Sharing in Business Units	territorialism	occasional data sharing	pockets of data sharing	managed data sharing	widespread data sharing
Data Sharing in Information Systems	individual interfaces	data sharing standards	shared databases	shared databases & metadata	data sharing architecture
Data Integration Technology	none	limited	adequate	good	excellent
Data Quality Technology	none	limited	adequate	good	excellent

TOTAL SCORE \_\_\_\_\_