SoberIT
Software Business and Engineering Institute

MDM and Data Governance

T-86.5161

Janne J. Korhonen

Helsinki University of Technology

Lecture Contents

- Master Data Management, lecture (40 min)
- SOA Characteristics and MDM, group work (60 min)
- Break (5 min)
- Data Governance, lecture (40 min)
- Review of Data Governance article (30 min)

MASTER DATA MANAGEMENT

Master Data

- *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.
  – Berson & Dubov (2007)

Master Data Management

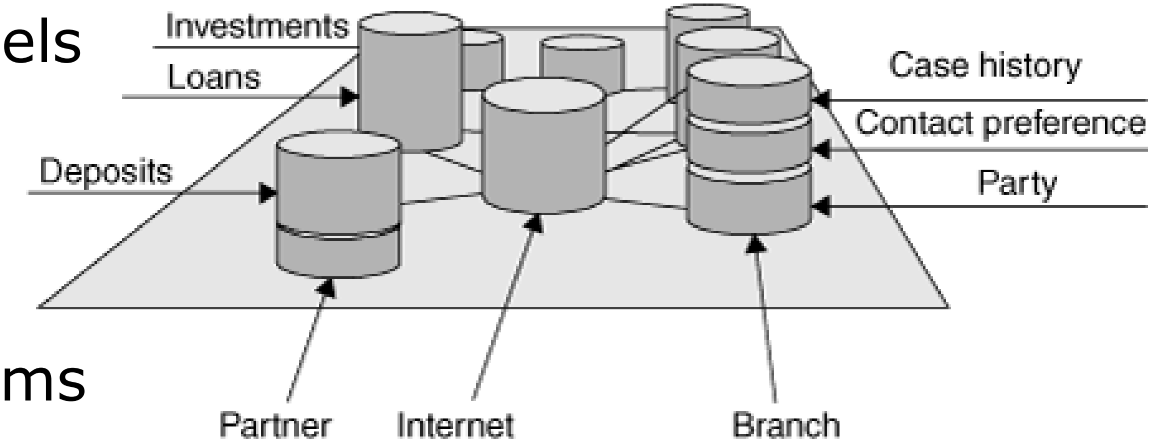
- *Master Data Management (MDM)* is the framework of processes and technologies aimed at creating and maintaining an authoritative, reliable, sustainable, accurate, and secure data environment that represents a "single version of truth," an accepted system of record used both intra- and inter-enterprise across a diverse set of application systems, lines of business, and user communities.
  – Berson & Dubov (2007)

MDM System

- Provides mechanisms for consistent use of master data across the organization
- Provides a consistent understanding and trust of master data entities
- Is designed to accommodate and manage change

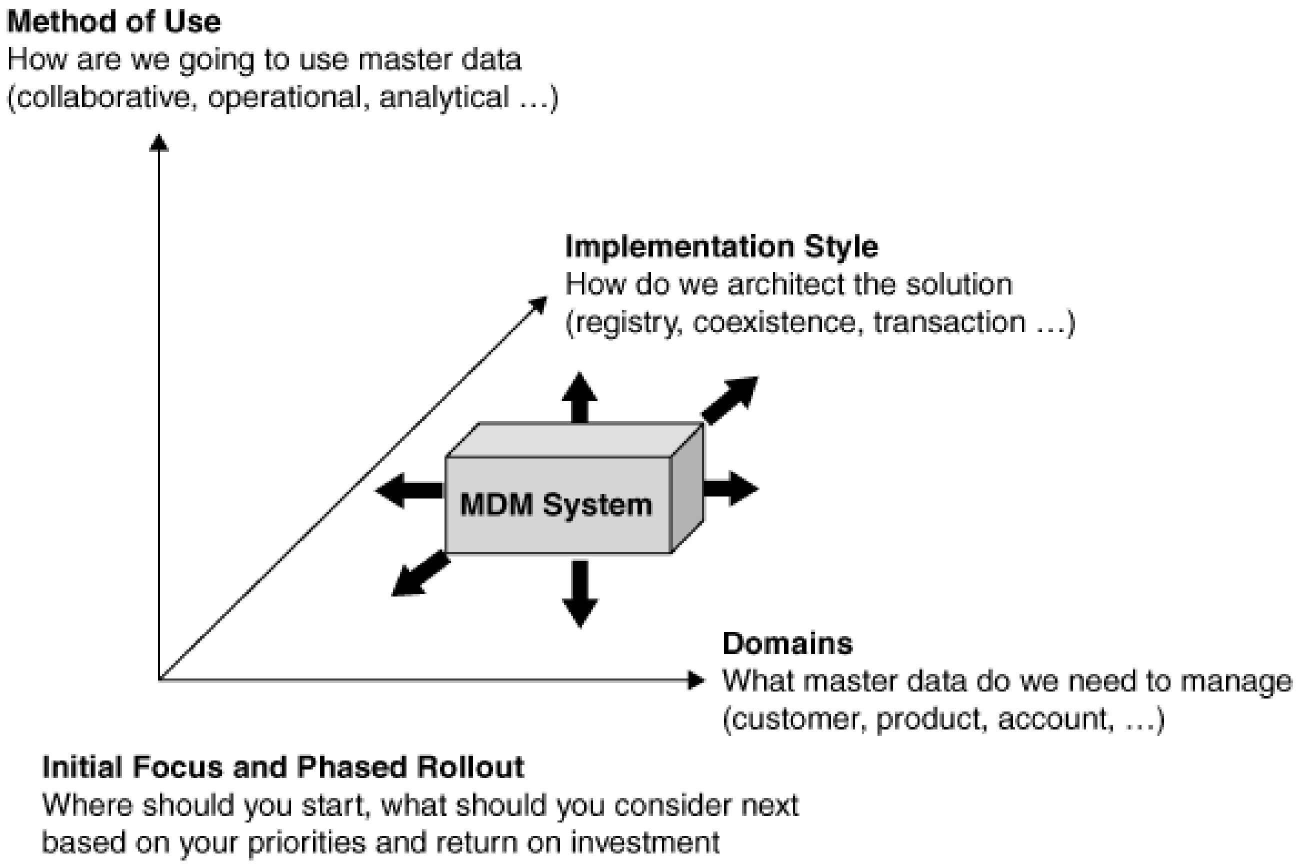
Why do organizations have multiple, often inconsistent, repositories of data?

- Line of business division
- Different channels
- Cross-domain distribution of information
- Packaged systems
- Mergers and acquisitions



Source: Dreibelbis et al (2008)

Dimensions of Master Data Management



Method of Use
How are we going to use master data
(collaborative, operational, analytical …)

Implementation Style
How do we architect the solution
(registry, coexistence, transaction …)

MDM System

Domains
What master data do we need to manage
(customer, product, account, …)

Initial Focus and Phased Rollout
Where should you start, what should you consider next
based on your priorities and return on investment

Source: Dreibelbis et al (2008)

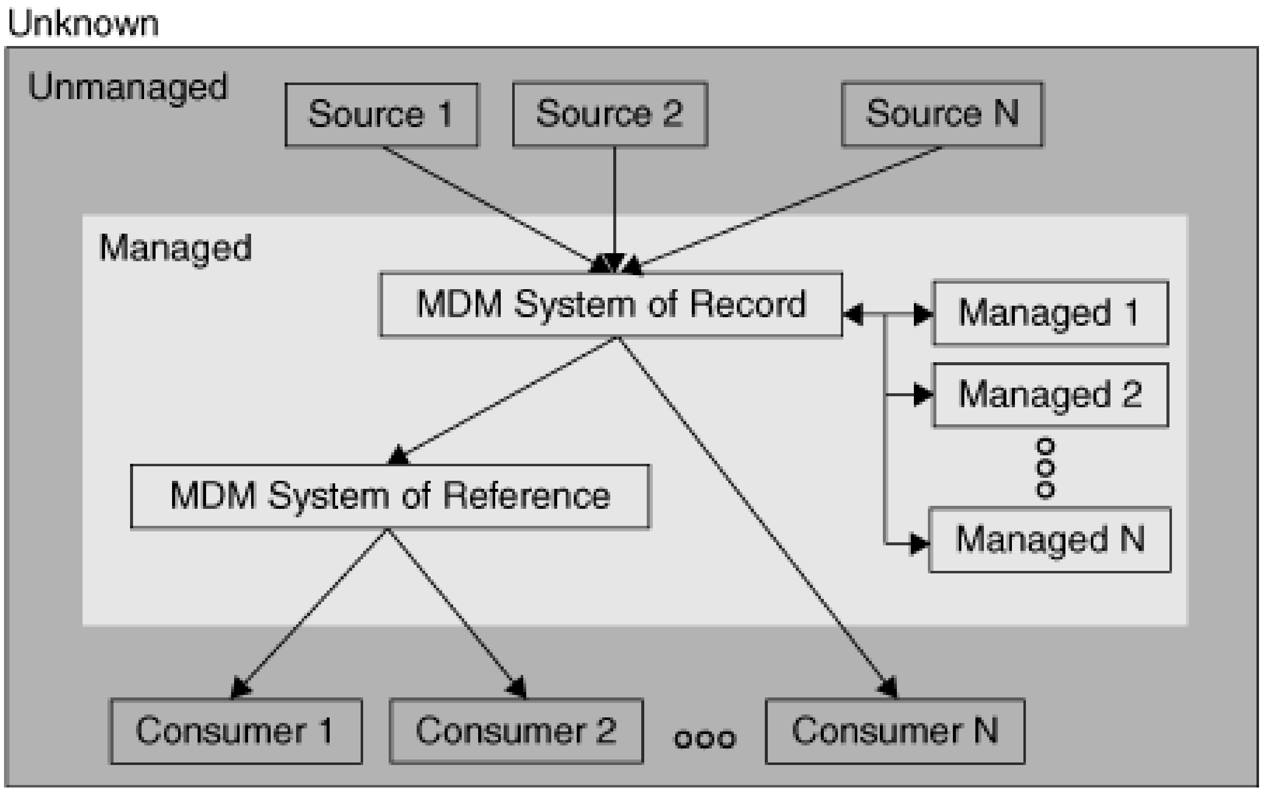
Master Data Domains

- Customer Data Integration (CDI)
- Product Information Management (PIM)
- Other domains: Accounts, Location...
- Industry Models:
  - Banking: Interactive Financial eXchange (IFX)
  - Telecom: Shared Information/Data Model (SID)
  - Healthcare: Health Level 7 (HL7)

Methods of Use

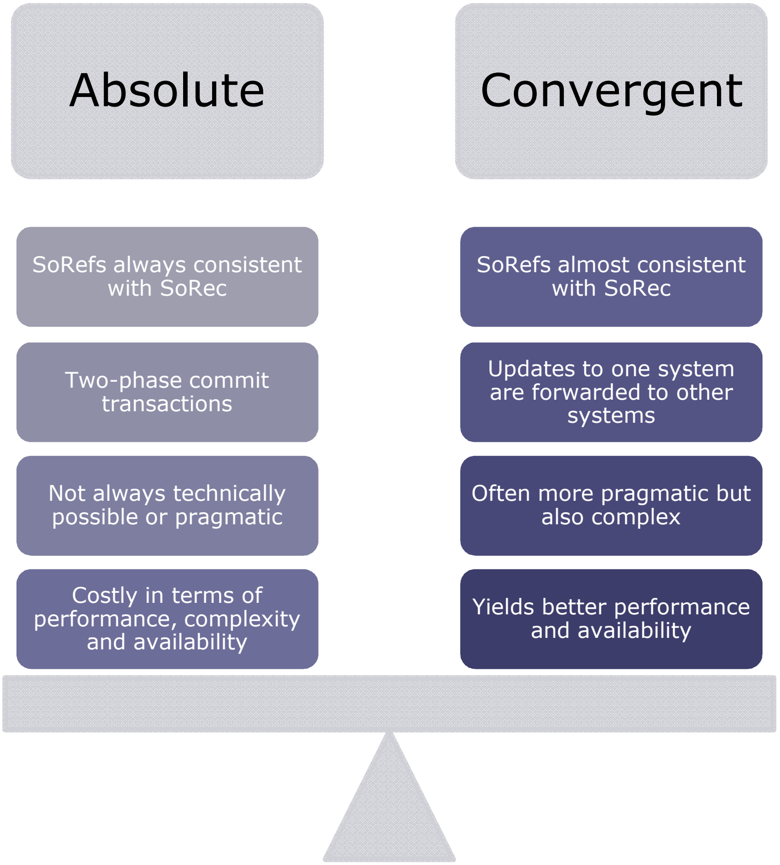
- Collaborative Authoring
  - MDM System coordinates a group of users and systems in order to reach agreement on a set of master data.
- Operational
  - MDM System participates in the operational transactions and business processes of the enterprise, interacting with other application systems and people.
- Analytical
  - MDM System is a source of authoritative information for downstream analytical systems, and sometimes is a source of insight itself.

System of Record vs. System of Reference



Source: Dreibelbis et al (2008)

Absolute vs. Convergent Consistency



Absolute

SoRefs always consistent with SoRec

Two-phase commit transactions

Not always technically possible or pragmatic

Costly in terms of performance, complexity and availability

Convergent

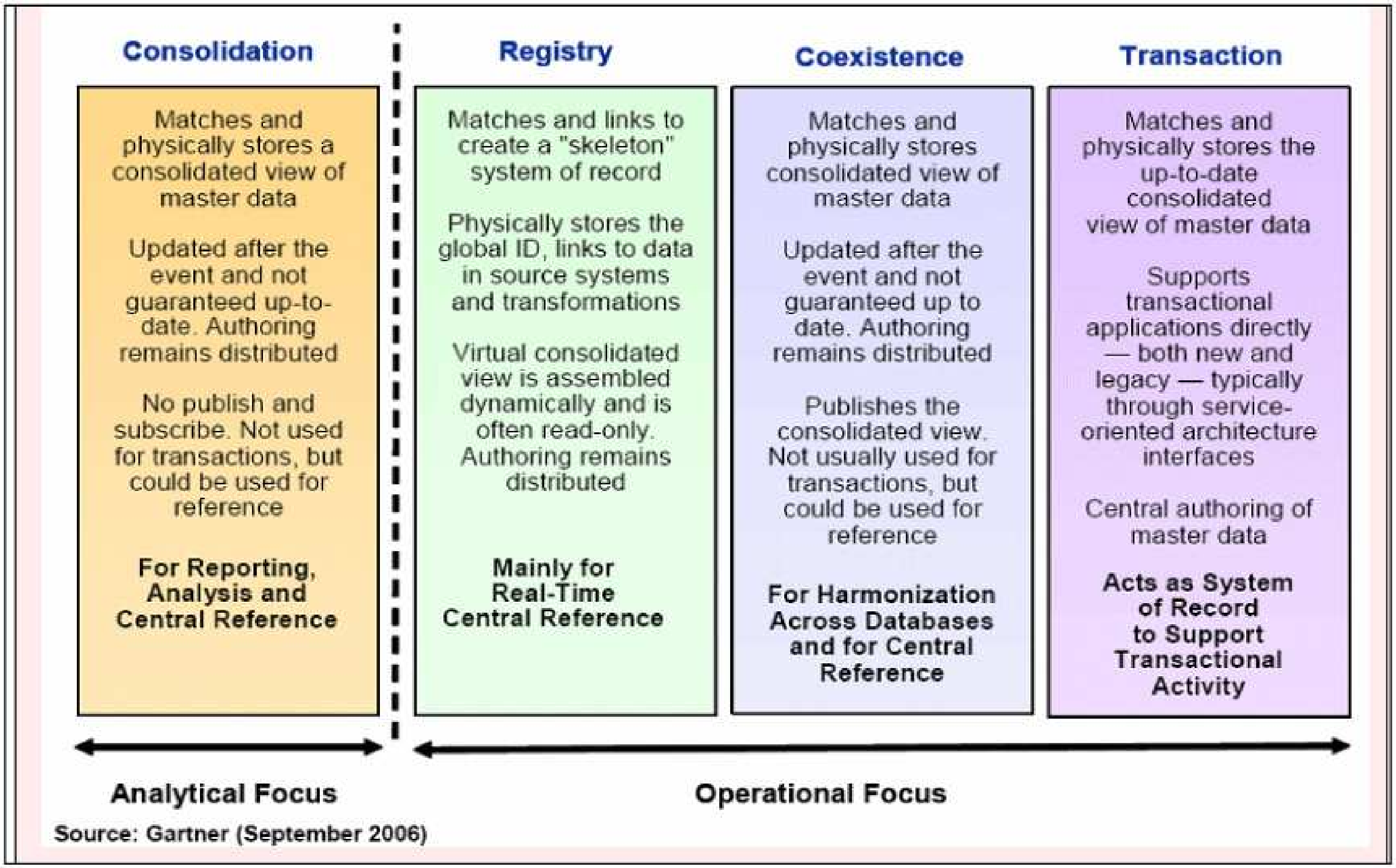
SoRefs almost consistent with SoRec

Updates to one system are forwarded to other systems

Often more pragmatic but also complex

Yields better performance and availability

MDM Implementation Styles

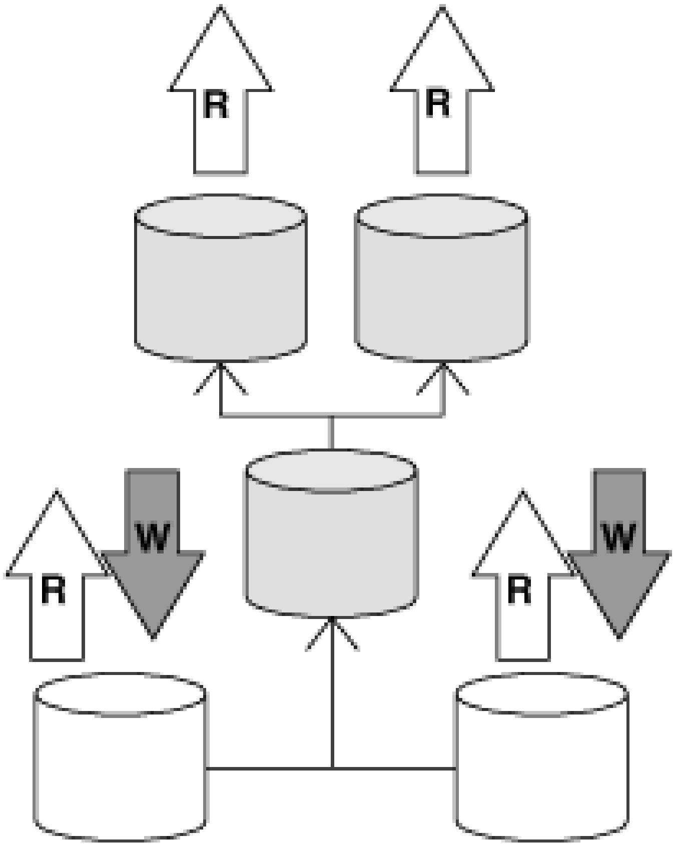


Consolidation	Registry	Coexistence	Transaction
Matches and physically stores a consolidated view of master data	Matches and links to create a "skeleton" system of record	Matches and physically stores consolidated view of master data	Matches and physically stores the up-to-date consolidated view of master data
Updated after the event and not guaranteed up-to-date. Authoring remains distributed	Physically stores the global ID, links to data in source systems and transformations	Updated after the event and not guaranteed up to date. Authoring remains distributed	Supports transactional applications directly — both new and legacy — typically through service-oriented architecture interfaces
No publish and subscribe. Not used for transactions, but could be used for reference	Virtual consolidated view is assembled dynamically and is often read-only. Authoring remains distributed	Publishes the consolidated view. Not usually used for transactions, but could be used for reference	Central authoring of master data
For Reporting, Analysis and Central Reference	**Mainly for Real-Time Central Reference**	**For Harmonization Across Databases and for Central Reference**	**Acts as System of Record to Support Transactional Activity**

Source: Gartner (September 2006)

Consolidation Implementation Style

- Brings together master data from a variety of existing systems into a single managed MDM hub
- The data is transformed, cleansed, matched, and integrated to provide a complete golden record for one or more master data domains
- A trusted source to downstream systems for reporting and analytics, or as a system of reference to other operational applications

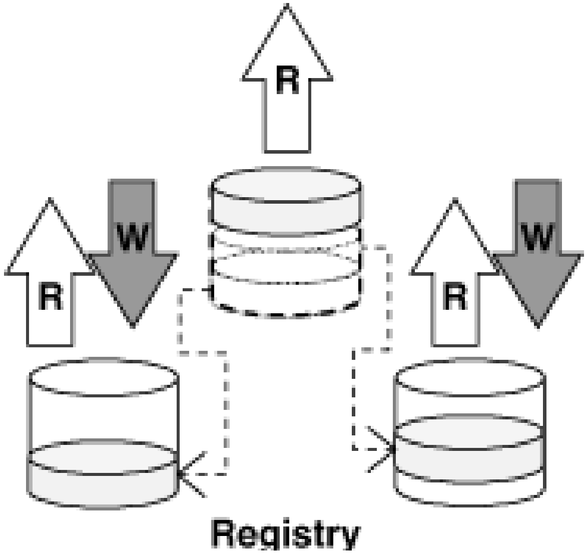


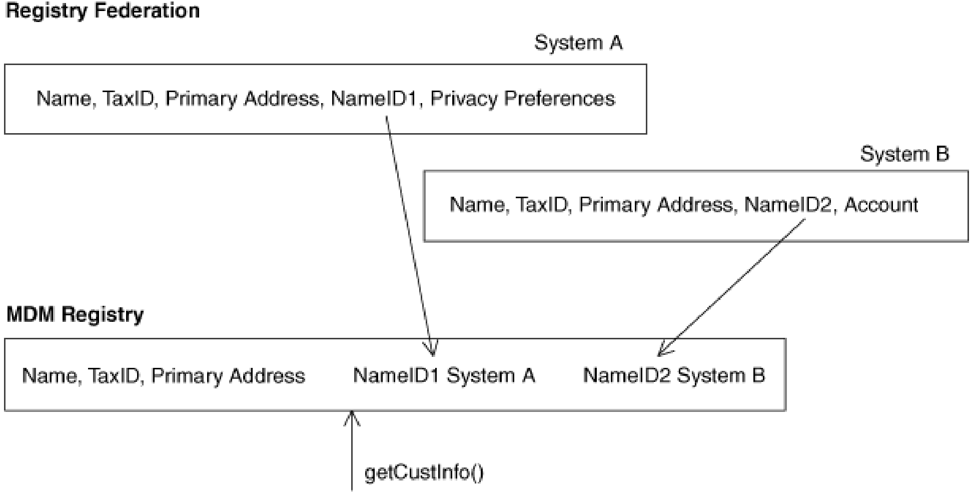
Consolidation

Registry Implementation Style

- Useful for providing a read-only source of master data as a reference to downstream systems with a minimum of data redundancy

- The registry is able to clean and match just the identifying cross reference information and assumes that the source systems are able to adequately manage the quality of their own data




Registry Federation

System A

Name, TaxID, Primary Address, NameID1, Privacy Preferences

System B

Name, TaxID, Primary Address, NameID2, Account

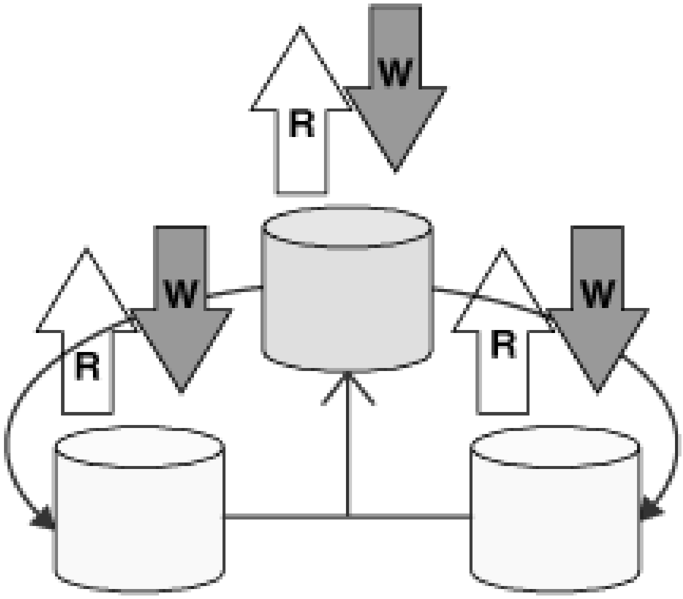
MDM Registry

Name, TaxID, Primary Address | NameID1 System A | NameID2 System B

getCustInfo()

Coexistence Implementation Style

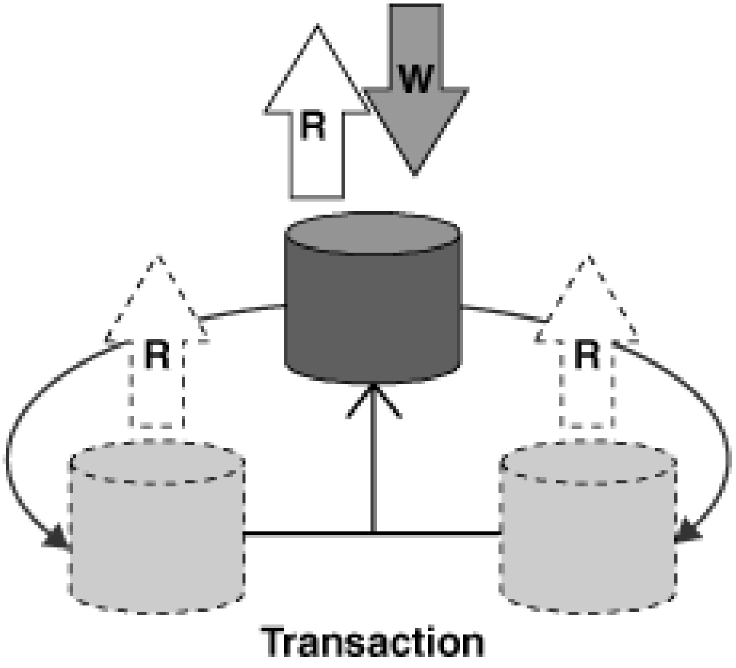
- Master data may be authored and stored in numerous locations
- Includes a physically instantiated golden record in the MDM System that is synchronized with source systems
- Not a system of record



Coexistence

Transactional Hub Implementation Style

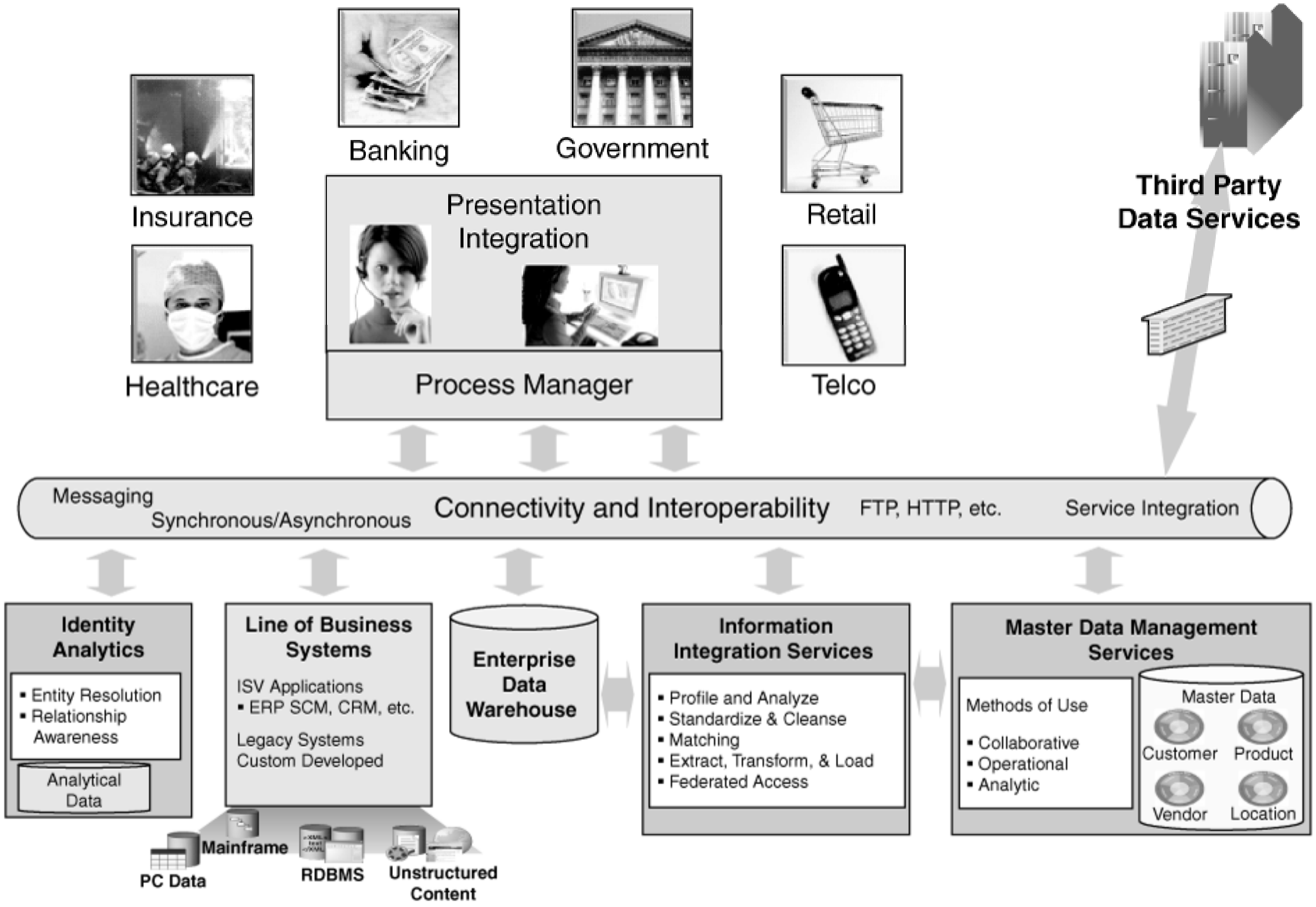
- A centralized, complete set of master data for one or more domains
- A system of record
- Often evolves from the consolidation and coexistence implementations



Comparison of Implementation Styles

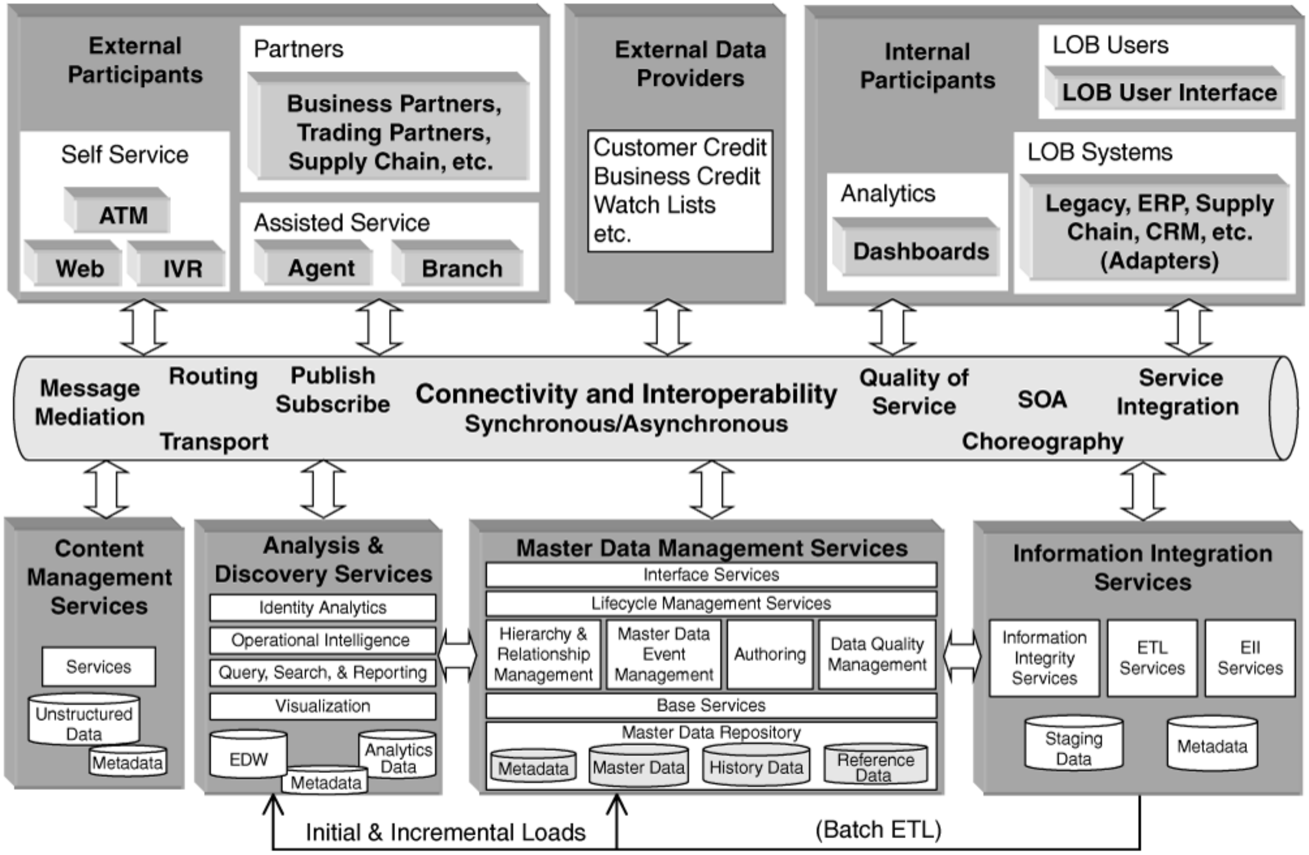
	Consolidation	Registry	Coexistence	Transactional Hub
What	Aggregate master data into a common repository for reporting and reference	Maintain thin system of record with links to more complete data spread across systems; useful for realtime reference	Manage single view of master data, synchronizing changes with other systems	Manage single view of master data, providing access via services
Benefits	Good for preparing data to feed downstream systems	Complete view is assembled as needed; fast to build	Assumes existing systems unchanged, yet provides read-write management	Support new and existing transactional applications; the system of record
Drawbacks	Read-only; not always current with operational systems	Read-mostly; may be more complex to manage	Not always consistent with other systems	May require changes to existing systems to exploit
Methods of use	Analytical	Operational	Collaborative, Operational, Analytical	Collaborative, Operational, Analytical
System of	Reference	Reference	Reference	Record

MDM Conceptual Architecture



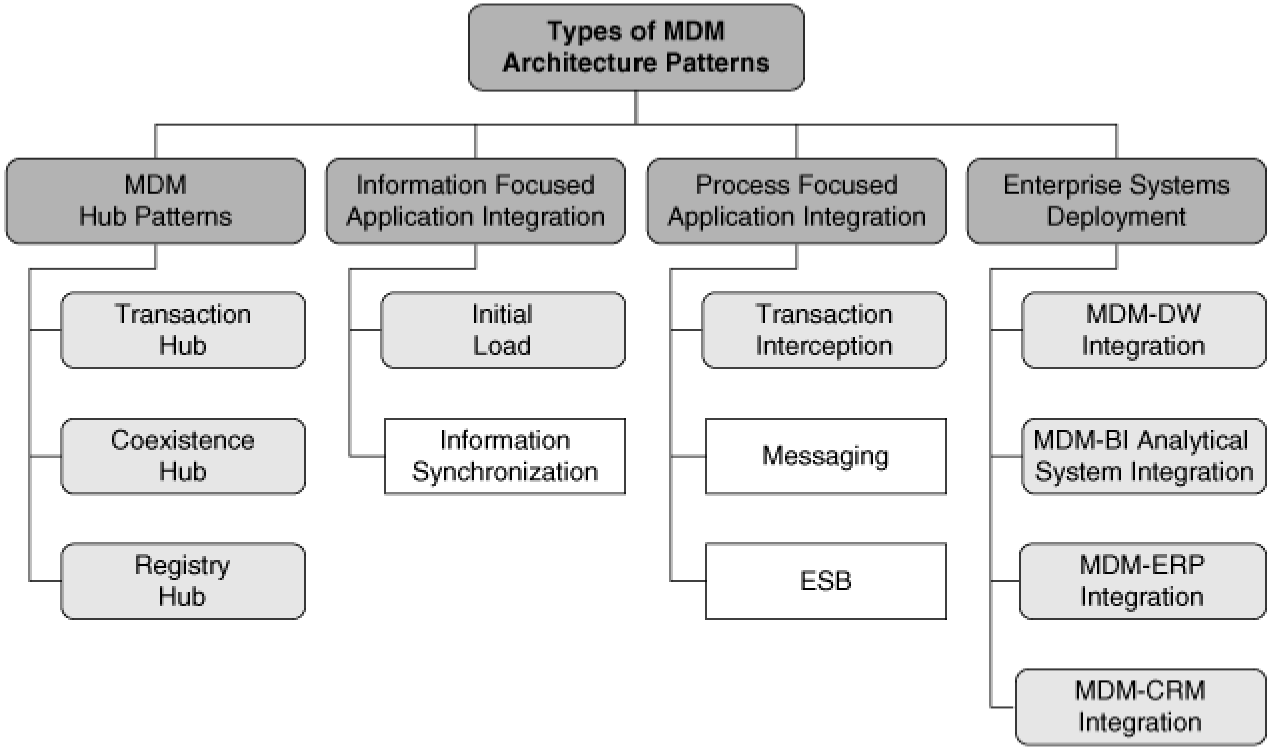
Source: Dreibelbis et al (2008)

MDM Logical Architecture



Source: Dreibelbis et al (2008)

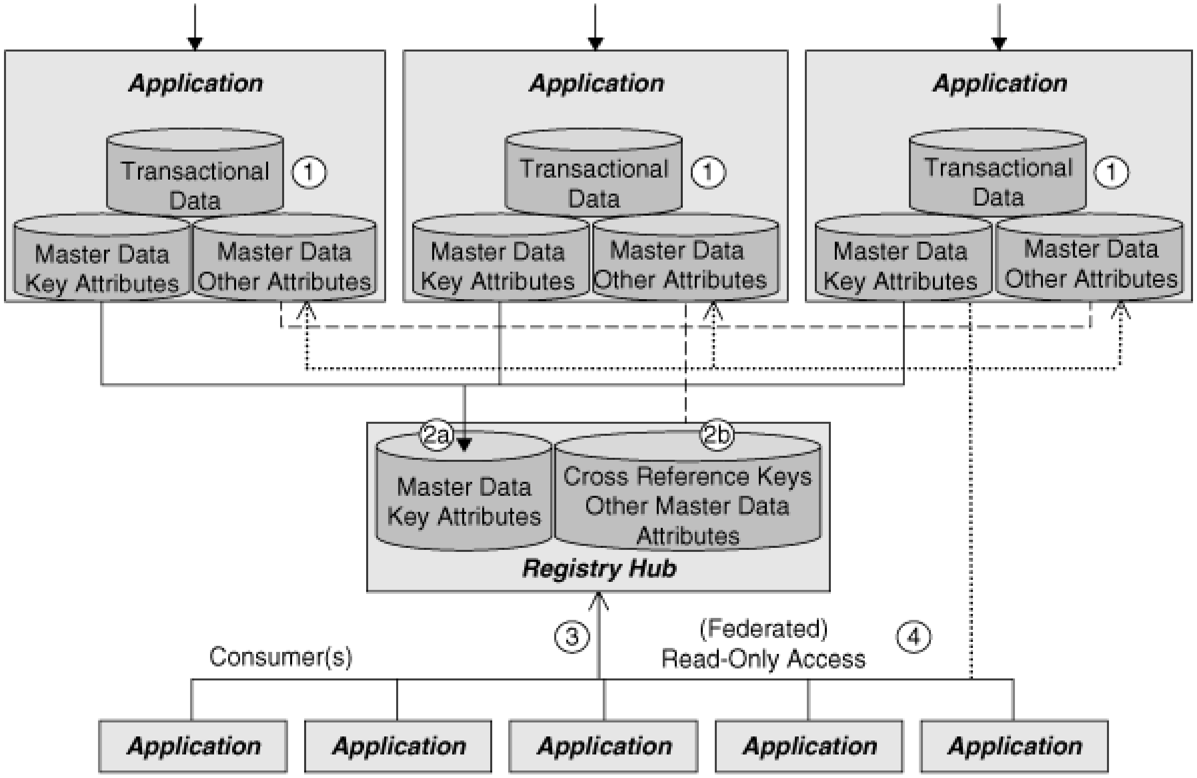
MDM Architecture Patterns


Types of MDM Architecture Patterns
MDM Hub Patterns
Information Focused Application Integration
Process Focused Application Integration
Enterprise Systems Deployment
Transaction Hub
Coexistence Hub
Registry Hub
Initial Load
Information Synchronization
Transaction Interception
Messaging
ESB
MDM-DW Integration
MDM-BI Analytical System Integration
MDM-ERP Integration
MDM-CRM Integration

Source: Dreibelbis et al (2008)

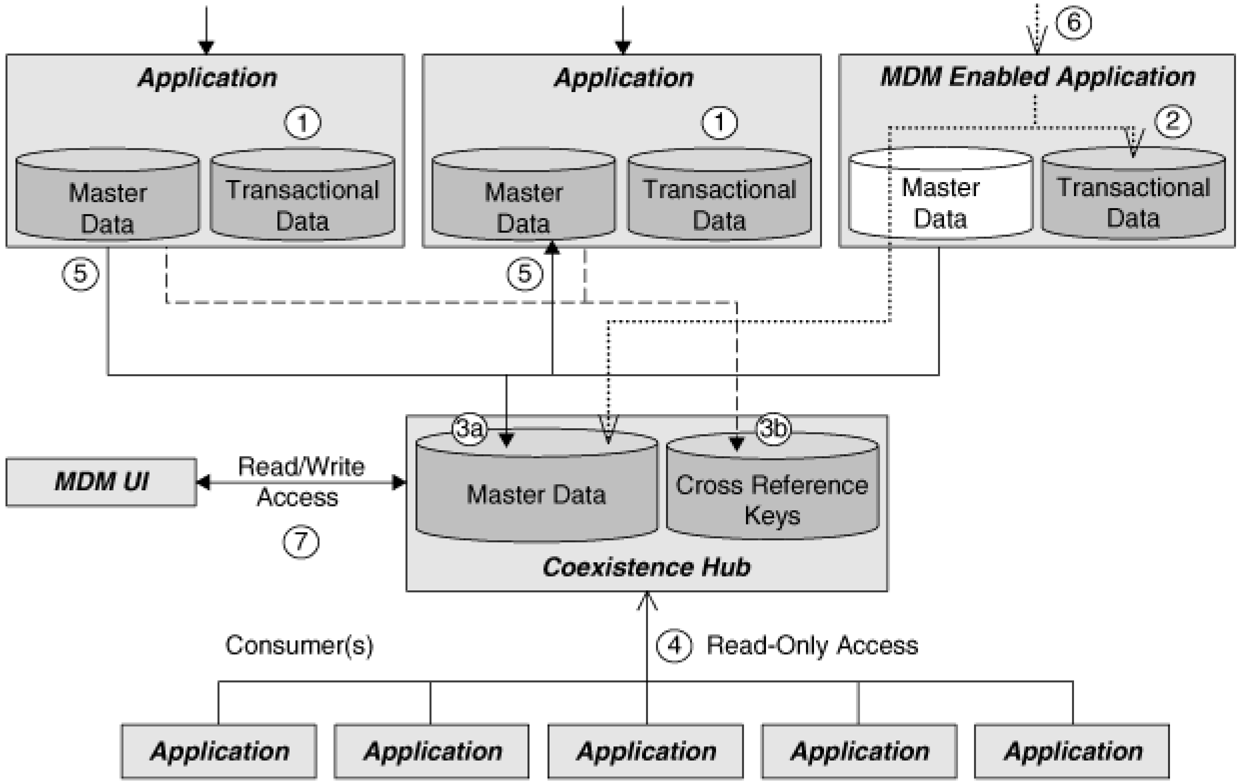
SoberIT
Software Business and Engineering Institute

Registry Hub Pattern


Application
Transactional Data
1
Master Data Key Attributes
Master Data Other Attributes
Application
Transactional Data
1
Master Data Key Attributes
Master Data Other Attributes
Application
Transactional Data
1
Master Data Key Attributes
Master Data Other Attributes
2a
Master Data Key Attributes
2b
Cross Reference Keys Other Master Data Attributes
Registry Hub
3
Consumer(s)
(Federated) Read-Only Access
4
Application
Application
Application
Application
Application

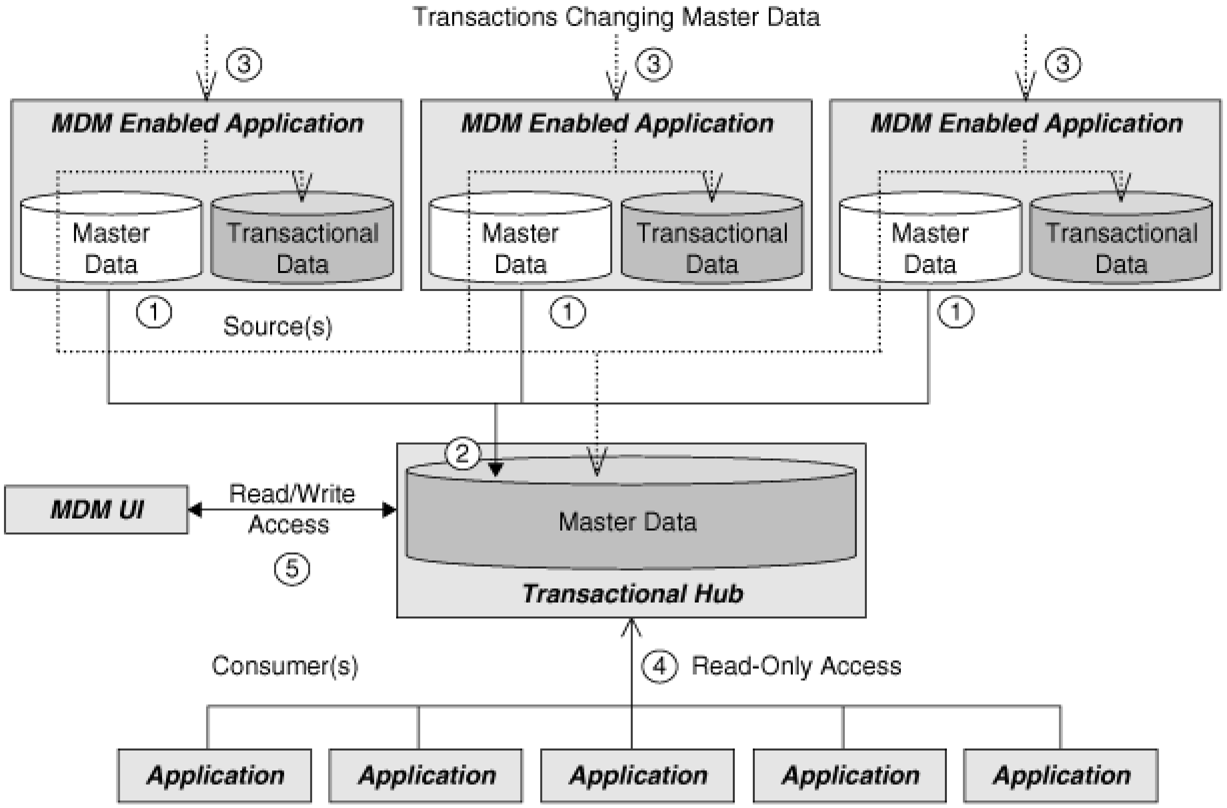
Source: Dreibelbis et al (2008)

Coexistence Hub Pattern



Source: Dreibelbis et al (2008)

Transaction Hub Patterm



Source: Dreibelbis et al (2008)

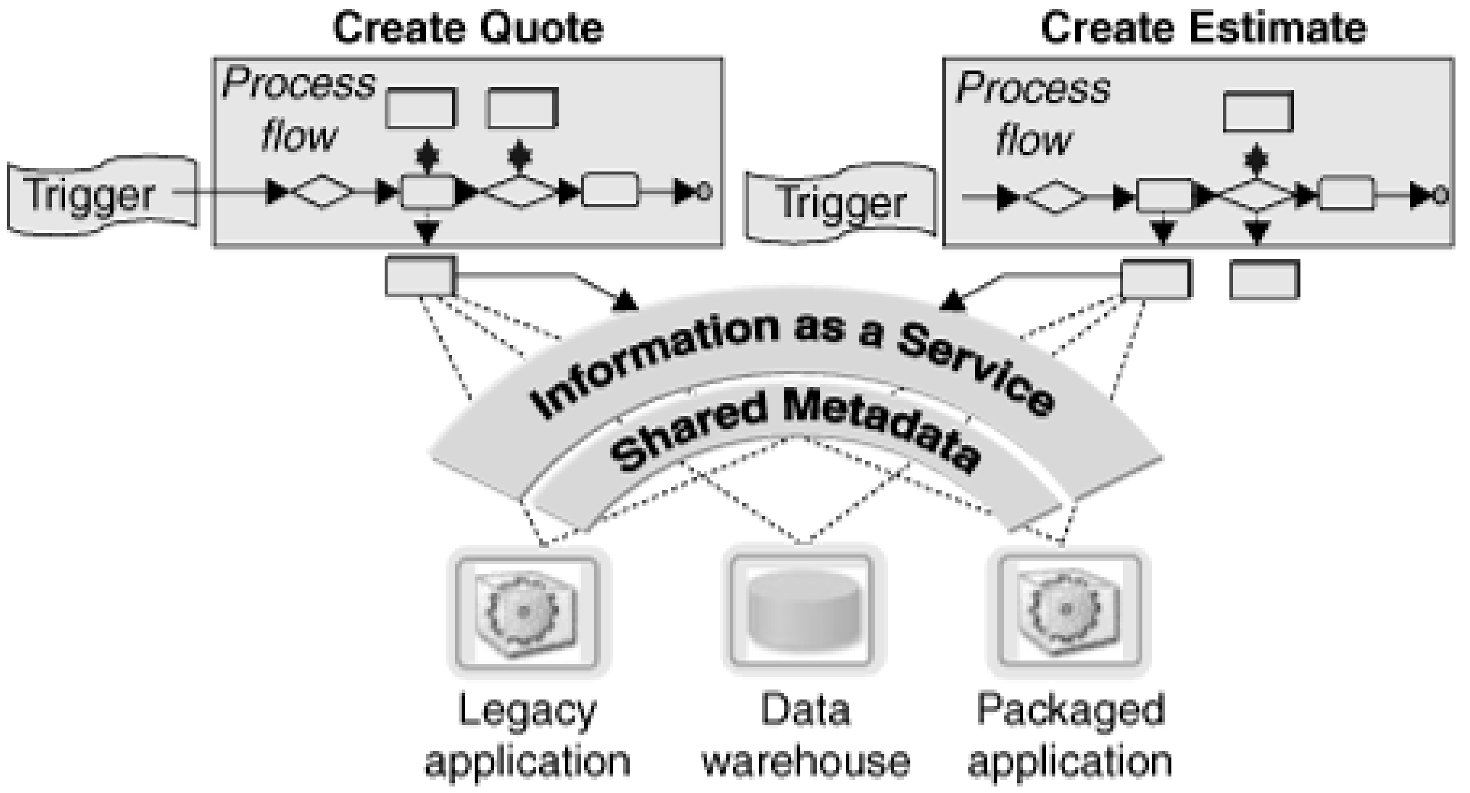
Comparison of MDM Hub Patterns

MDM Hub Pattern	Registry Hub	Coexistence Hub	Transaction Hub
Purpose	Central reference	Harmonization	Transactional access
System Type	System of reference	System of reference	System of record
Method of Use	(Analytical, collaborative), operational	Analytical, collaborative, operational	Analytical, collaborative, operational
Master Data	Stored in legacy systems, but can be consistently viewed and derived through linkage to master data in these legacy systems	Stored in MDM and legacy systems, where the MDM System serves as a base for a single source of truth	Stored in MDM and legacy systems, where the MDM System serves as a base for a single source of truth
Master Data Services	Master data creation and maintenance done in legacy systems	Master data creation and maintenance done in legacy systems and the MDM Hub as well	Master data creation and maintenance only done through MDM services provided by MDM System
Type of Access/Transactions	Read only, where insert, update, and delete statements can only be performed against the legacy systems	Read only, where some of the insert, update, and delete statements will be performed against the MDM Hub, and some of these statements will be performed against the legacy systems	Read and write, where insert, update, and delete statements can be performed directly against the master data in the hub

Comparison of MDM Hub Patterns

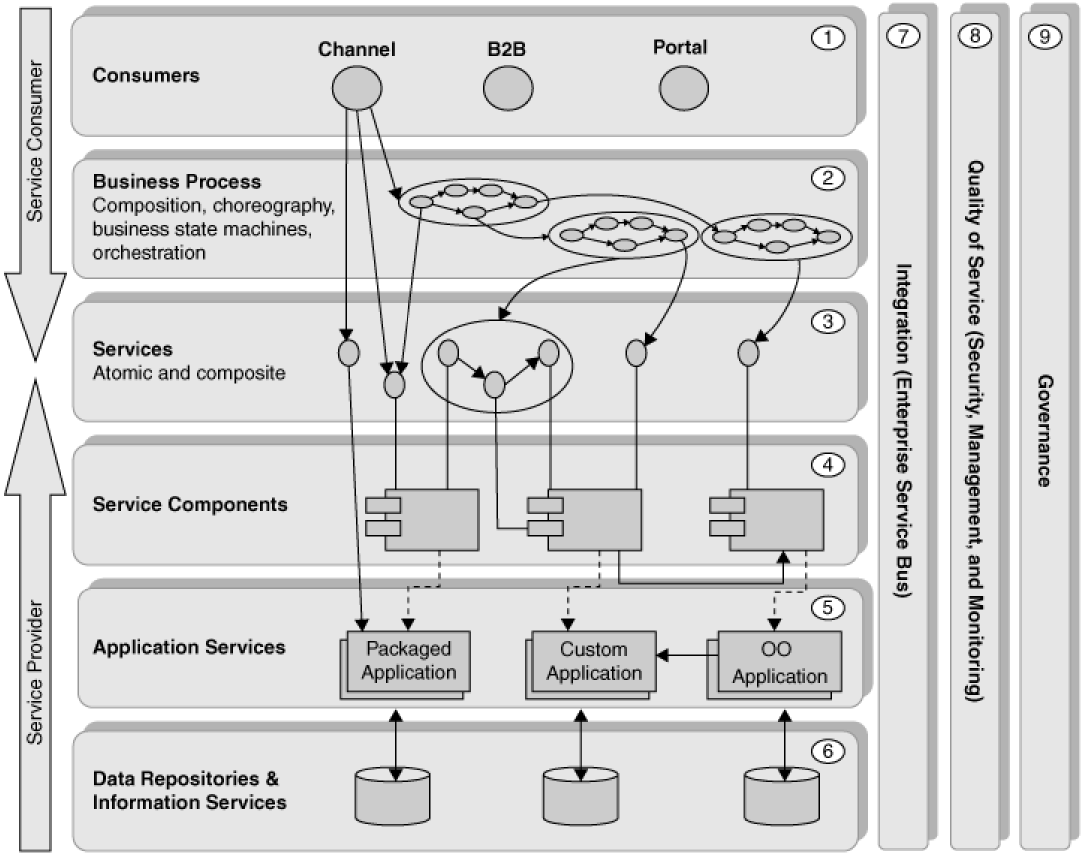
MDM Hub Pattern	Registry Hub	Coexistence Hub	Transaction Hub
Correctness	Only key attributes are materialized (cleansed, de-duplicated) in MDM System All other attributes remain unchanged (low quality) in legacy system	On initial load, all master data attributes are cleansed, standardized, and de-duplicated when materialized in MDM System On change, correctness delayed in MDM System due to potential delay in propagation	Given at all times, because access is through MDM services, which incorporate cleanse, standardize, and de-duplication routines
Completeness	Only through reference to legacy system achieved by virtualization/ federation	Complete, because fully materialized on initial load	Complete, because fully materialized on initial load
Consistency	No consistency: Master data remains inconsistent in legacy systems	Converging consistency: Multiple legacy and MDM System are updated; conflicts require resolution	Converging to absolute consistency: Transactions are invoked only through MDM services of MDM System and propagated to consuming applications (asynchronously or synchronously)

Information as a Service (IaaS)



Source: Dreibelbis et al (2008)

Service-Oriented Architecture



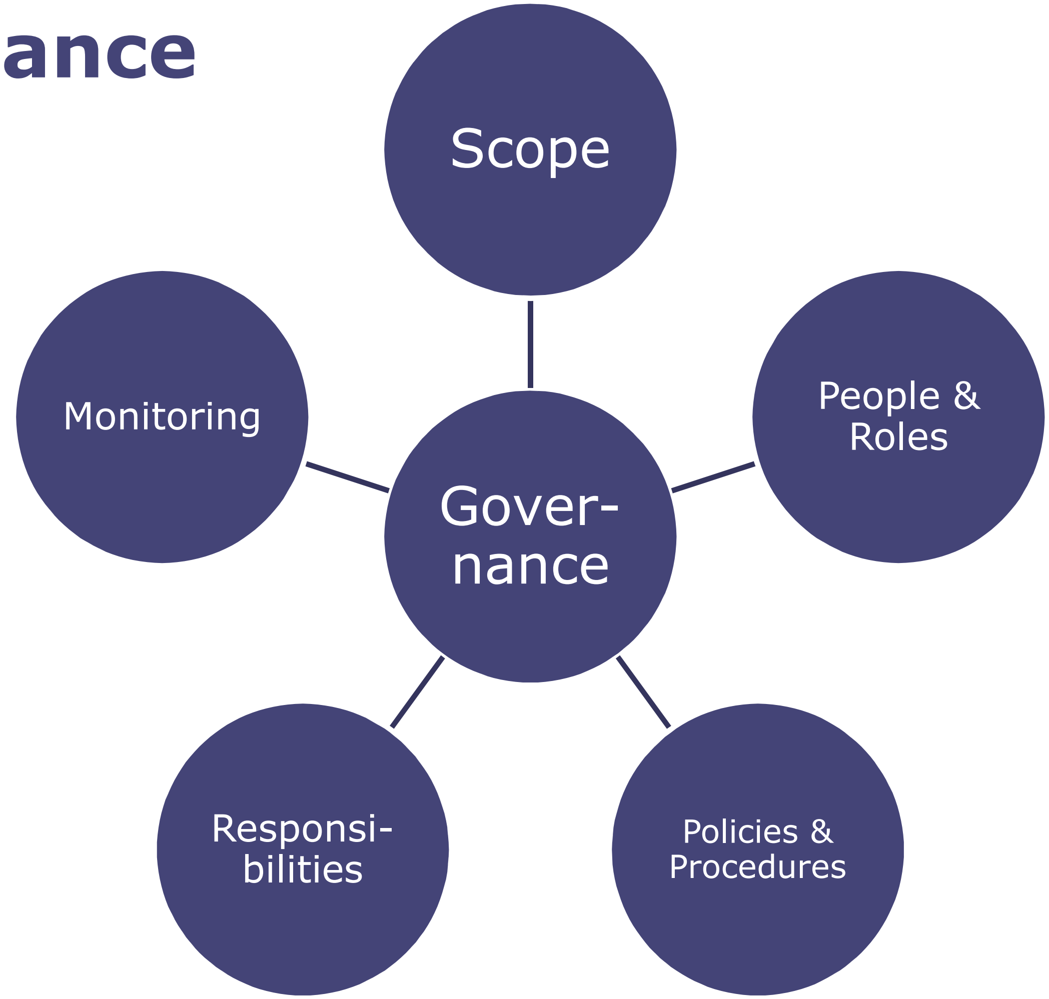
Source: Dreibelbis et al (2008)

SOA Characteristics and MDM

- **Service reuse**
- **Service granularity**
- **Service modularity and loose coupling**
- **Service composability**
- **Service componentization and encapsulation**
- Compliance with standards (both common and industry-specific)
- Services identification and categorization
- Provisioning and delivery
- Monitoring and tracking

DATA GOVERNANCE

Governance



Adapted from Dreibelbis et al (2008)

Governance in IT

- Mature IT organizations have a set of processes, policies, and procedures related to IT architecture, including:
    - Defining architectural components, behaviors, interfaces, and integration
    - Getting approval of the architecture
    - Ensuring that the IT infrastructure and applications align with architecture standards
    - Requesting changes to the components to accommodate new application requirements or emerging technologies
    - Granting variances to application architects and owners for all or part of the architectural requirements

Control Objectives for Information and related Technology (CobiT)

- Best practices (framework) for information technology (IT) management
- Created by the Information Systems Audit and Control Association (ISACA) and the IT Governance Institute (ITGI)
- Covers four domains:
  - Plan and Organize
  - Acquire and Implement
  - Deliver and Support
  - Monitor and Evaluate

CobiT Example:
DS11 Manage Data – Process Description

Control over the IT process of

Manage data

that satisfies the business requirement for IT of

optimising the use of information and ensuring that information is available as required

by focusing on

maintaining the completeness, accuracy, availability and protection of data

is achieved by

- Backing up data and testing restoration
- Managing onsite and offsite storage of data
- Securely disposing of data and equipment

and is measured by

- Percent of user satisfaction with availability of data
- Percent of successful data restorations
- Number of incidents where sensitive data were retrieved after media were disposed

DS11 Manage Data – Control Objectives

DS11.1 Business Requirements for Data Management

- Verify that all data expected for processing are received and processed completely, accurately and in a timely manner, and all output is delivered in accordance with business requirements. Support restart and reprocessing needs.

DS11.2 Storage and Retention Arrangements

- Define and implement procedures for effective and efficient data storage, retention and archiving to meet business objectives, the organisation's security policy and regulatory requirements.

DS11.3 Media Library Management System

- Define and implement procedures to maintain an inventory of stored and archived media to ensure their usability and integrity.

DS11.4 Disposal

- Define and implement procedures to ensure that business requirements for protection of sensitive data and software are met when data and hardware are disposed or transferred.

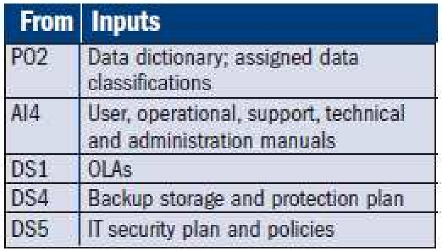
DS11.5 Backup and Restoration

- Define and implement procedures for backup and restoration of systems, applications, data and documentation in line with business requirements and the continuity plan.

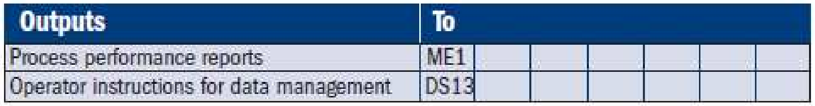
DS11.6 Security Requirements for Data Management

- Define and implement policies and procedures to identify and apply security requirements applicable to the receipt, processing, storage and output of data to meet business objectives, the organisation's security policy and regulatory requirements.

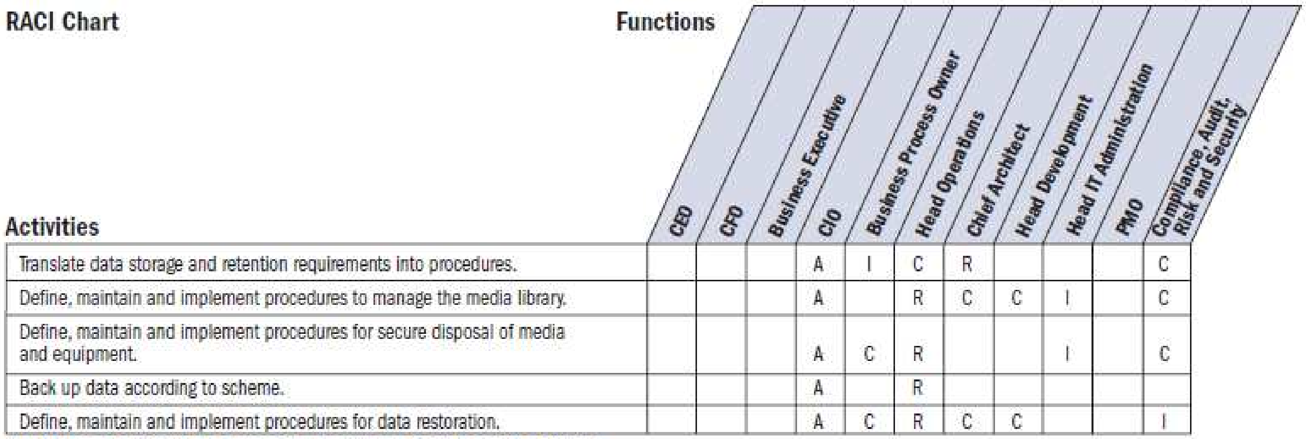
DS11 Manage Data – Management Guidelines



From	Inputs
PO2	Data dictionary; assigned data classifications
AI4	User, operational, support, technical and administration manuals
DS1	OLAs
DS4	Backup storage and protection plan
DS5	IT security plan and policies



Outputs	To						
Process performance reports	ME1						
Operator instructions for data management	DS13						

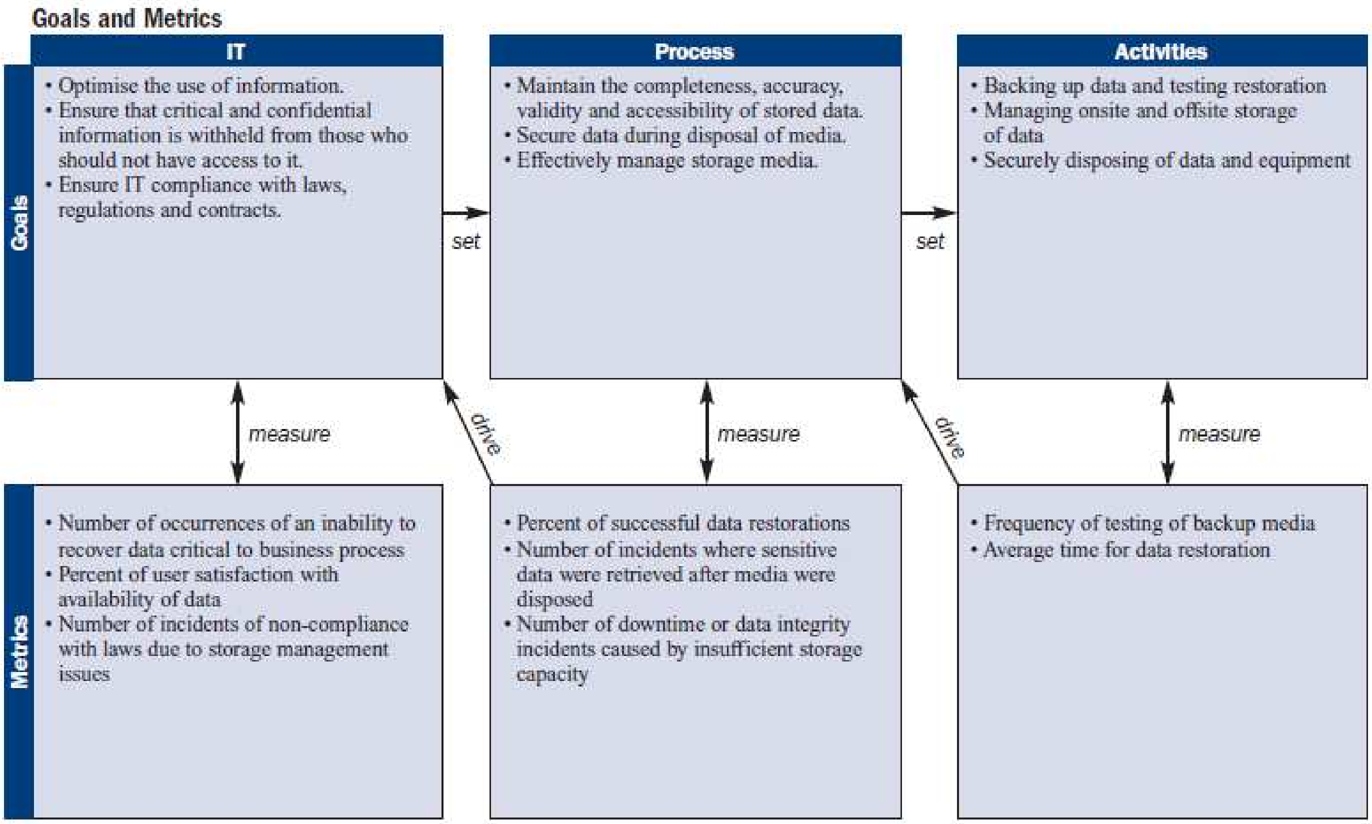


RACI Chart

Activities	CEO	CFO	Business Executive	CIO	Business Process Owner	Head Operations	Chief Architect	Head Development	Head IT Administration	PMO	Compliance, Audit, Risk and Security
Translate data storage and retention requirements into procedures.				A	I	C	R				C
Define, maintain and implement procedures to manage the media library.				A		R	C	C	I		C
Define, maintain and implement procedures for secure disposal of media and equipment.				A	C	R			I		C
Back up data according to scheme.				A		R					
Define, maintain and implement procedures for data restoration.				A	C	R	C	C			I

A **RACI** chart identifies who is **R**esponsible, **A**ccountable, **C**onsulted and/or **I**nformed.

DS11 Manage Data – Management Guidelines



Goals and Metrics

Goals

IT

- Optimise the use of information.
- Ensure that critical and confidential information is withheld from those who should not have access to it.
- Ensure IT compliance with laws, regulations and contracts.

Process

- Maintain the completeness, accuracy, validity and accessibility of stored data.
- Secure data during disposal of media.
- Effectively manage storage media.

Activities

- Backing up data and testing restoration
- Managing onsite and offsite storage of data
- Securely disposing of data and equipment

Metrics

- Number of occurrences of an inability to recover data critical to business process
- Percent of user satisfaction with availability of data
- Number of incidents of non-compliance with laws due to storage management issues

- Percent of successful data restorations
- Number of incidents where sensitive data were retrieved after media were disposed
- Number of downtime or data integrity incidents caused by insufficient storage capacity

- Frequency of testing of backup media
- Average time for data restoration

DS11 Manage Data – Maturity Model

Management of the process of *Manage data that satisfies the business requirement for IT of optimising the use of information and ensuring that information is available as required is:*

0 Non-existent when

Data are not recognised as corporate resources and assets. There is no assigned data ownership or individual accountability for data management. Data quality and security are poor or non-existent.

1 Initial/Ad Hoc when

The organisation recognises a need for effective data management. There is an *ad hoc approach for specifying security requirements* for data management, but no formal communications procedures are in place. No specific training on data management takes place. Responsibility for data management is not clear. Backup/restoration procedures and disposal arrangements are in place.

2 Repeatable but Intuitive when

The awareness of the need for effective data management exists throughout the organisation. Data ownership at a high level begins to occur. Security requirements for data management are documented by key individuals. Some monitoring within IT is performed on data management key activities (e.g., backup, restoration, disposal). Responsibilities for data management are informally assigned for key IT staff members.

DS11 Manage Data – Maturity Model

3 Defined when

The need for data management within IT and across the organisation is understood and accepted. Responsibility for data management is established. Data ownership is assigned to the responsible party who controls integrity and security. Data management procedures are formalised within IT, and some tools for backup/restoration and disposal of equipment are used. Some monitoring over data management is in place. Basic performance metrics are defined. Training for data management staff members is emerging.

4 Managed and Measurable when

The need for data management is understood, and required actions are accepted within the organisation. Responsibility for data ownership and management are clearly defined, assigned and communicated within the organisation. Procedures are formalised and widely known, and knowledge is shared. Usage of current tools is emerging. Goal and performance indicators are agreed to with customers and monitored through a well-defined process. Formal training for data management staff members is in place.

5 Optimised when

The need for data management and the understanding of all required actions is understood and accepted within the organisation. Future needs and requirements are explored in a proactive manner. The responsibilities for data ownership and data management are clearly established, widely known across the organisation and updated on a timely basis. Procedures are formalised and widely known, and knowledge sharing is standard practice. Sophisticated tools are used with maximum automation of data management. Goal and performance indicators are agreed to with customers, linked to business objectives and consistently monitored using a well-defined process. Opportunities for improvement are constantly explored. Training for data management staff members is instituted.

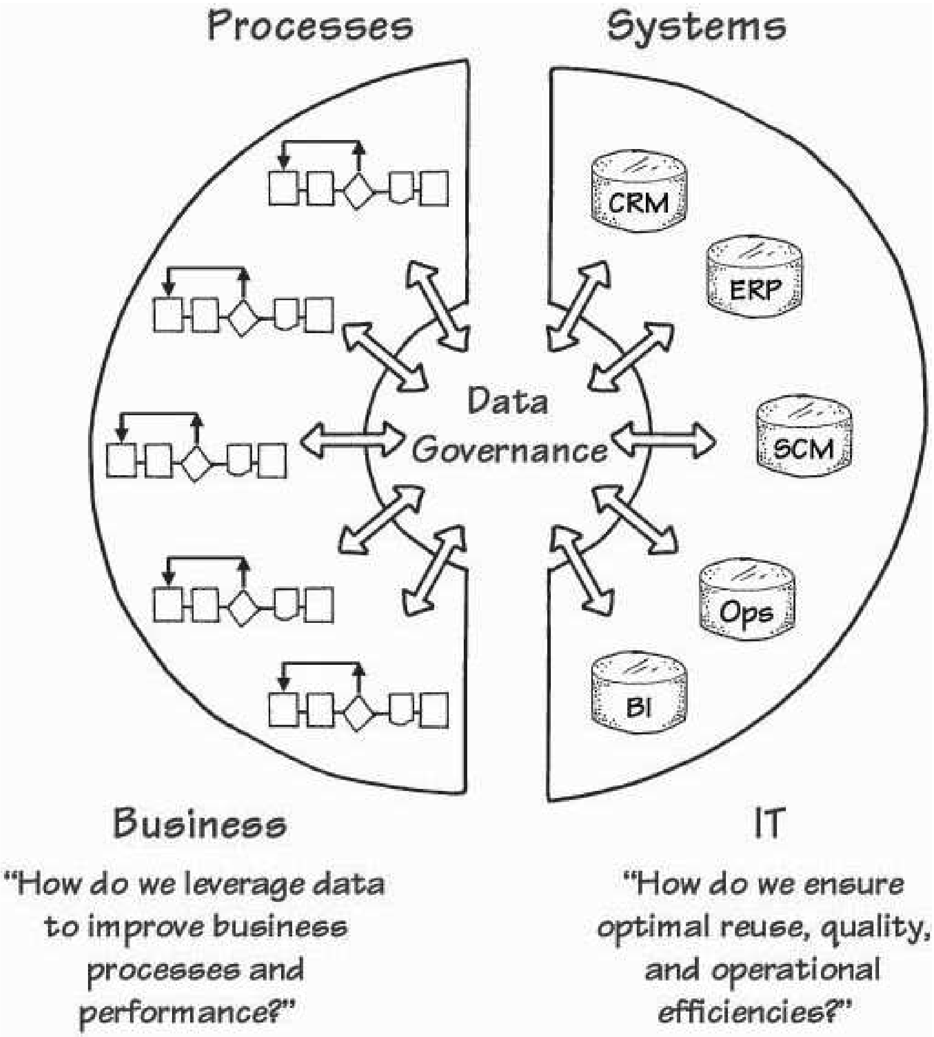
IT Governance and Data Governance

- IT is like the pipes and pumps and storage tanks in a plumbing system
- Data is like the water flowing through those pipes
- If you suspected your water was poisoned, would you call a plumber?

Data Governance Defined

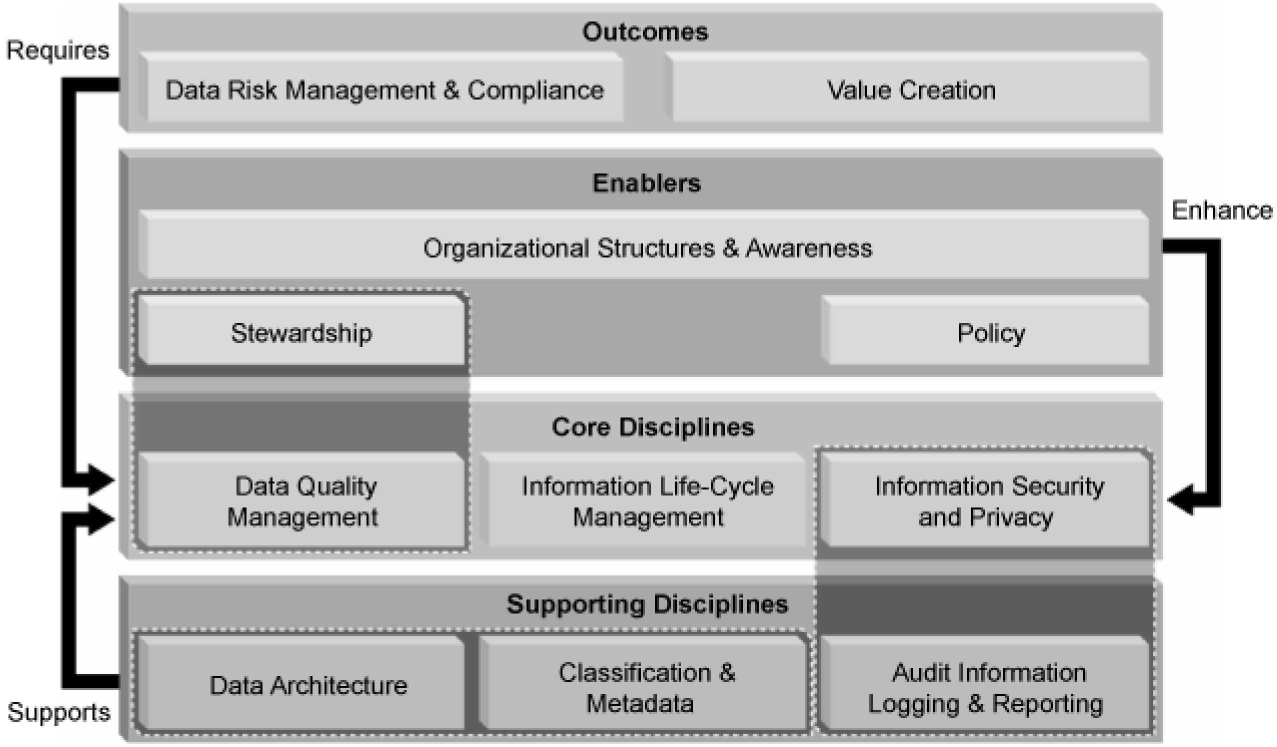
- As defined by the IBM Data Governance Council, data governance is
  - *"the political process of changing organizational behavior to enhance and protect data as a strategic enterprise asset."*
- The Data Governance Institute defines DG as
  - *"a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods."*

Data Governance Touches Both Business and IT

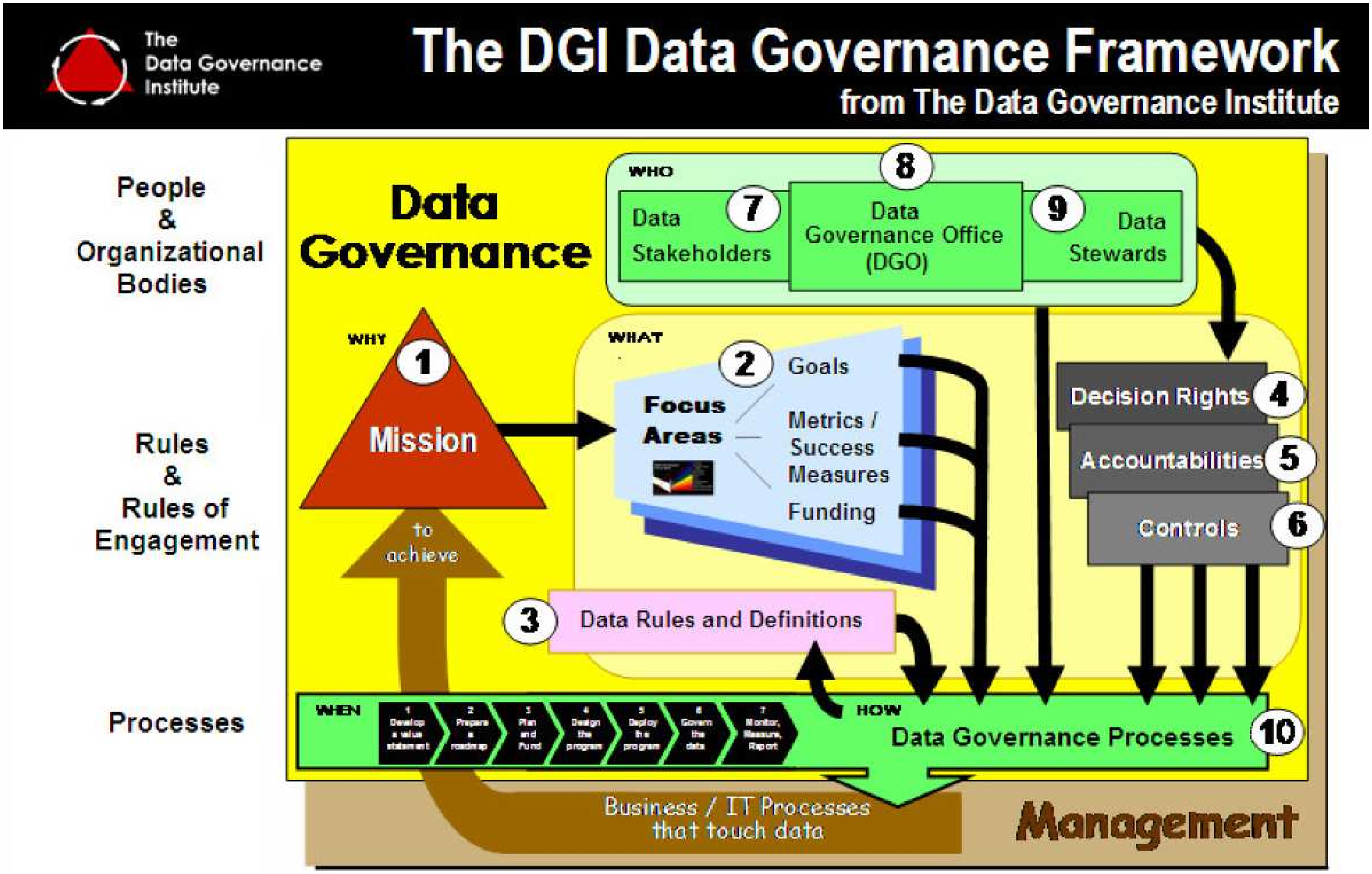

Processes
Systems
CRM
ERP
Data
Governance
SCM
Ops
BI
Business
IT
"How do we leverage data to improve business processes and performance?"
"How do we ensure optimal reuse, quality, and operational efficiencies?"

Source: Dyché & Levy (2006)

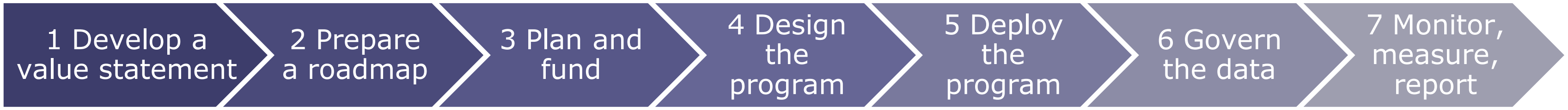
Disciplines of Effective Data Governance


Requires
Outcomes
Data Risk Management & Compliance
Value Creation
Enablers
Enhance
Organizational Structures & Awareness
Stewardship
Policy
Core Disciplines
Data Quality Management
Information Life-Cycle Management
Information Security and Privacy
Supporting Disciplines
Data Architecture
Classification & Metadata
Audit Information Logging & Reporting
Supports

Source: the IBM Data Governance Council


The Data Governance Institute
The DGI Data Governance Framework
from The Data Governance Institute
People & Organizational Bodies
Rules & Rules of Engagement
Processes
Data Governance
WHO
8
Data Stakeholders
7
Data Governance Office (DGO)
9
Data Stewards
WHY
1
Mission
to achieve
WHAT
2
Focus Areas
Goals
Metrics / Success Measures
Funding
Decision Rights
4
Accountabilities
5
Controls
6
3
Data Rules and Definitions
WHEN
1 Develop a value statement
2 Prepare a roadmap
3 Plan and Fund
4 Design the program
5 Deploy the program
6 Govern the data
7 Monitor, Measure, Report
HOW
Data Governance Processes
10
Business / IT Processes that touch data
Management

Data Governance Program Lifecycle



1 Develop a value statement

2 Prepare a roadmap

3 Plan and fund

4 Design the program

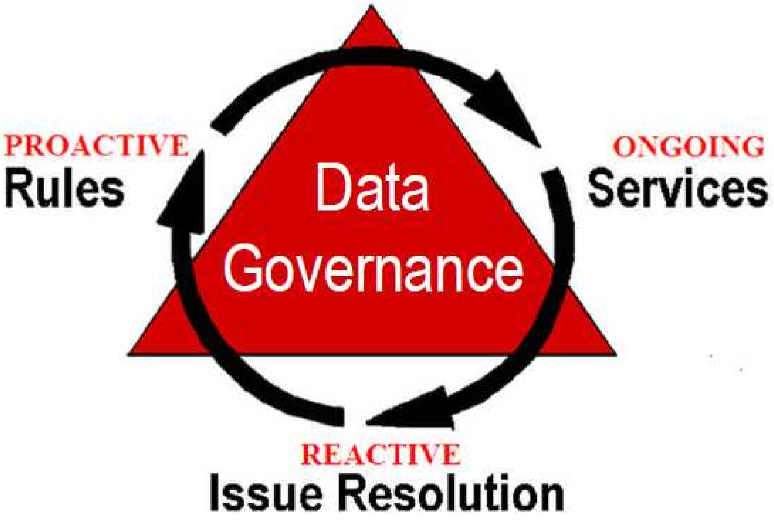
5 Deploy the program

6 Govern the data

7 Monitor, measure, report

Data Governance typically has a three-part mission

1. Proactively define/align rules
2. Provide ongoing, boundary-spanning protection and services to data stakeholders
3. React to and resolve issues arising from non-compliance with rules



Typical Universal Goals of a Data Governance Program

1. Enable better decision-making
2. Reduce operational friction
3. Protect the needs of data stakeholders
4. Train management and staff to adopt common approaches to data issues
5. Build standard, repeatable processes
6. Reduce costs and increase effectiveness through coordination of efforts
7. Ensure transparency of processes

Data Governance Focus Areas

- Policy, Standards, Strategy
- Data Quality
- Privacy / Compliance / Security
- Architecture / Integration
- Data Warehouses and BI
- Management Alignment

Data-Related Rules

- Policies
- Standards
- Guidelines
- Requirements
- Guiding Principles
- Business Rules

- Data Quality Rules
- Data Usage Rules
- Data Access Rules
- “Golden Copy” designations
- “System of Record” designations
- etc.

Ha ha!

GOVERNANCE IN ACTION
Most of our Data Governance policies are more specific than "Because I said so."
www.ComicStripGenerator.com

Considerations

Current State

- How does the organization deal with policies, standards, and other types of rules?
- Who can create them?
- Where are they stored?
- How are they disseminated?
- Do managers enforce them?
- How do business and technical staff respond to them?
- What other groups work with data-related rules?
- What type of alignment is missing between these groups?

Future Vision

- What type of alignment should you aim for with these groups?
- What are your organization's pain points, and how might Data Governance address them?
- What types of rules would need to be in formalized to do this?
- What data subject areas should the rules be initially applied to? What about later?
- What areas of the data environment would initially be affected? What about later?
- What compliance rules will need to be considered?
- Will it be more effective to introduce a group of rules all at once, or a few at a time?
- Who needs to approve rules that come from the Data Governance program? Who should be consulted before they are finalized? Who should be informed before they are announced?

Post-Compliance Paradigm

- For efforts with a compliance requirement, the work is not finished until we

1. Do the work
2. Control it
3. Document it, and then
4. Prove compliance.

Source: http://www.tdan.com/view-special-features/5356

A Model For Data Governance (Wende 2007)

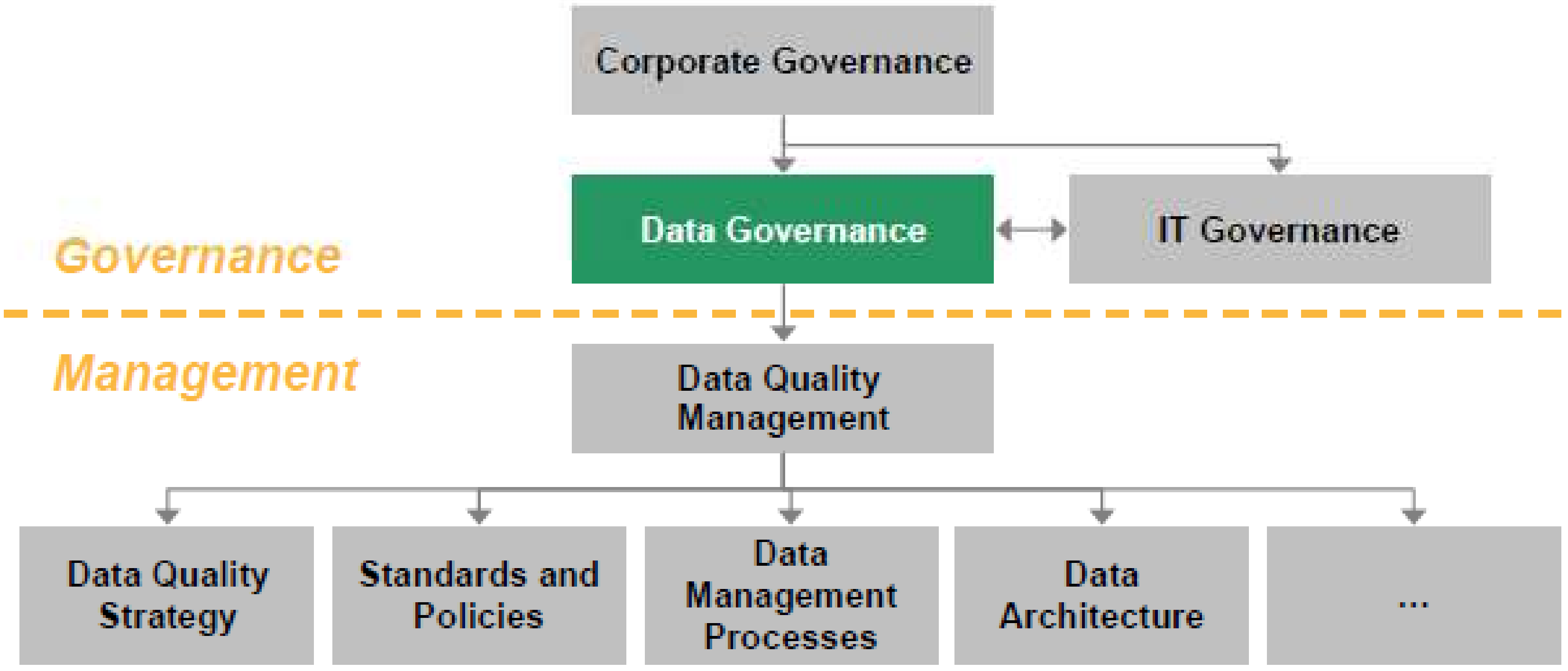
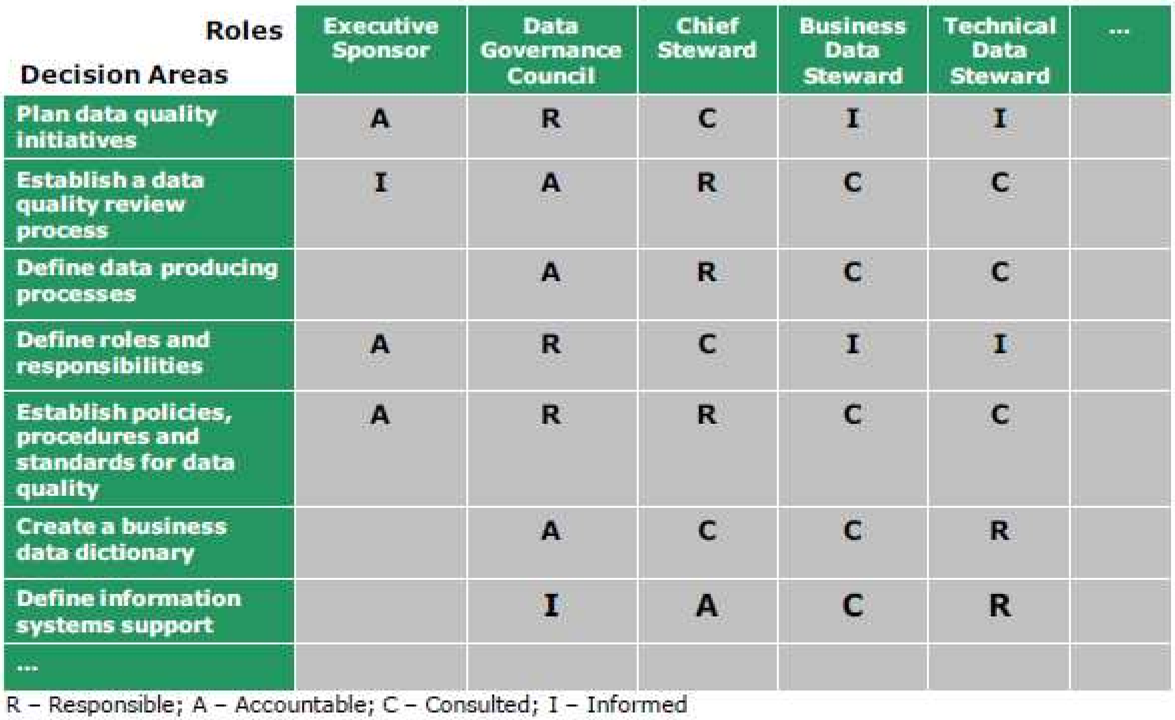


Figure 1: Terms in Governance and Management

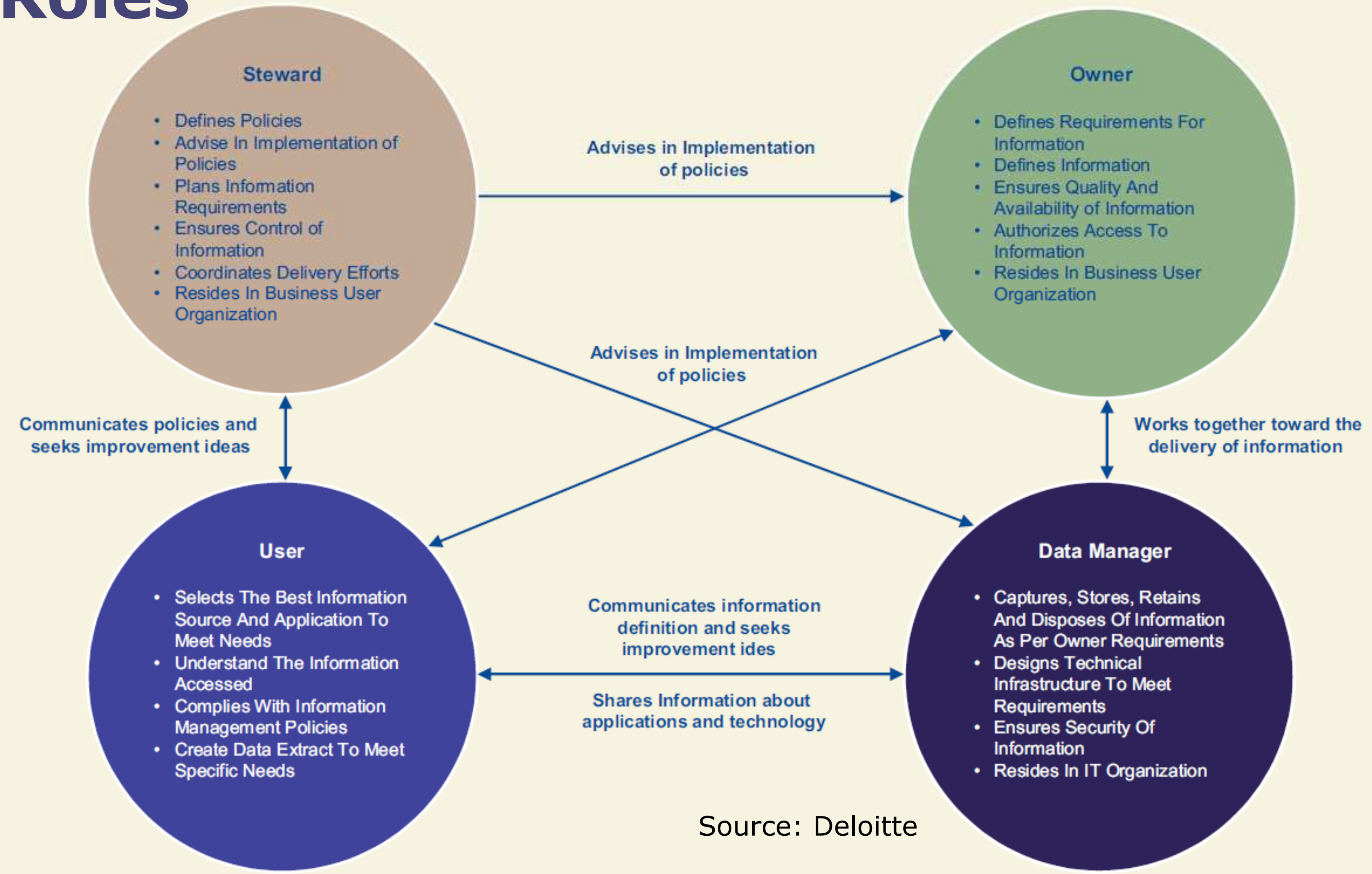
Data Governance Model (Wende 2007)



Roles Decision Areas	Executive Sponsor	Data Governance Council	Chief Steward	Business Data Steward	Technical Data Steward	...
Plan data quality initiatives	A	R	C	I	I	
Establish a data quality review process	I	A	R	C	C	
Define data producing processes		A	R	C	C	
Define roles and responsibilities	A	R	C	I	I	
Establish policies, procedures and standards for data quality	A	R	R	C	C	
Create a business data dictionary		A	C	C	R	
Define information systems support		I	A	C	R	
...						

R – Responsible; A – Accountable; C – Consulted; I – Informed

Typical Data Governance Organization Roles



Steward

- Defines Policies
- Advise In Implementation of Policies
- Plans Information Requirements
- Ensures Control of Information
- Coordinates Delivery Efforts
- Resides In Business User Organization

Owner

- Defines Requirements For Information
- Defines Information
- Ensures Quality And Availability of Information
- Authorizes Access To Information
- Resides In Business User Organization

User

- Selects The Best Information Source And Application To Meet Needs
- Understand The Information Accessed
- Complies With Information Management Policies
- Create Data Extract To Meet Specific Needs

Data Manager

- Captures, Stores, Retains And Disposes Of Information As Per Owner Requirements
- Designs Technical Infrastructure To Meet Requirements
- Ensures Security Of Information
- Resides In IT Organization

Advises in Implementation of policies

Advises in Implementation of policies

Communicates policies and seeks improvement ideas

Works together toward the delivery of information

Communicates information definition and seeks improvement ides

Shares Information about applications and technology

Source: Deloitte

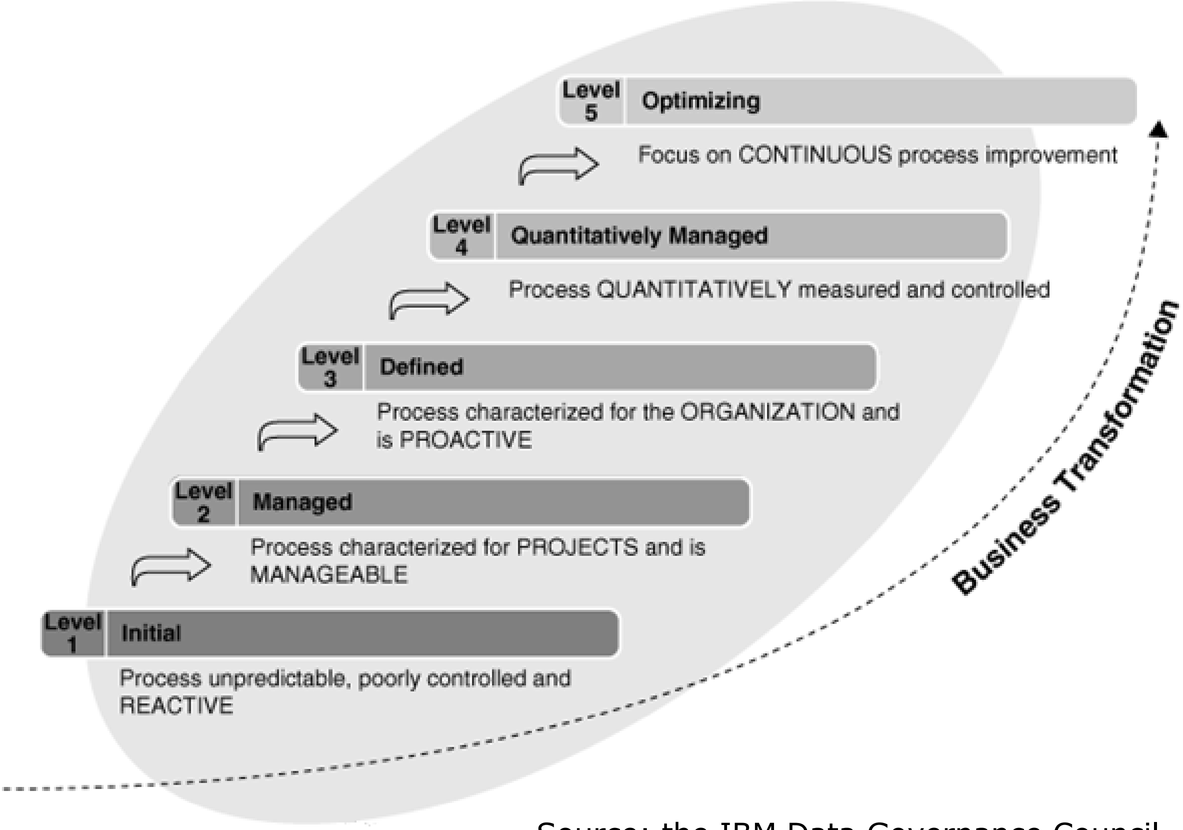
Common Responsibilities of Data Governance Office (DGO)

- run the program
- keep track of Data Stakeholders and Stewards
- serve as liaison to other discipline and programs, such a Data Quality, Compliance, Privacy, Security, Architecture, and IT Governance
- collect and align policies, standards, and guideline from these stakeholder group
- arrange for the providing of information and analysis to IT projects as requested
- facilitate and coordinate meetings of Data Stewards
- collect metric and success measures and report on them to data stakeholders
- provide ongoing Stakeholder CARE in the form of **C**ommunication, **A**ccess to information, **R**ecordkeeping, and **E**ducation/support
- articulate the value of Data Governance and Stewardship activities
- provide centralized communication for governance-led and data-related matters

Common Data Governance Processes

1. Aligning Policies, Requirements, and Controls
2. Establishing Decision Rights
3. Establishing Accountability
4. Performing Stewardship
5. Managing Change
6. Defining Data
7. Resolving Issues
8. Specifying Data Quality Requirements
9. Building Governance Into Technology
10. Providing Stakeholder Care
11. Communications and Program Reporting.
12. Measuring and Reporting Value

Data Governance Maturity Model (DGMM)



Level 5 Optimizing

Focus on CONTINUOUS process improvement

Level 4 Quantitatively Managed

Process QUANTITATIVELY measured and controlled

Level 3 Defined

Process characterized for the ORGANIZATION and is PROACTIVE

Level 2 Managed

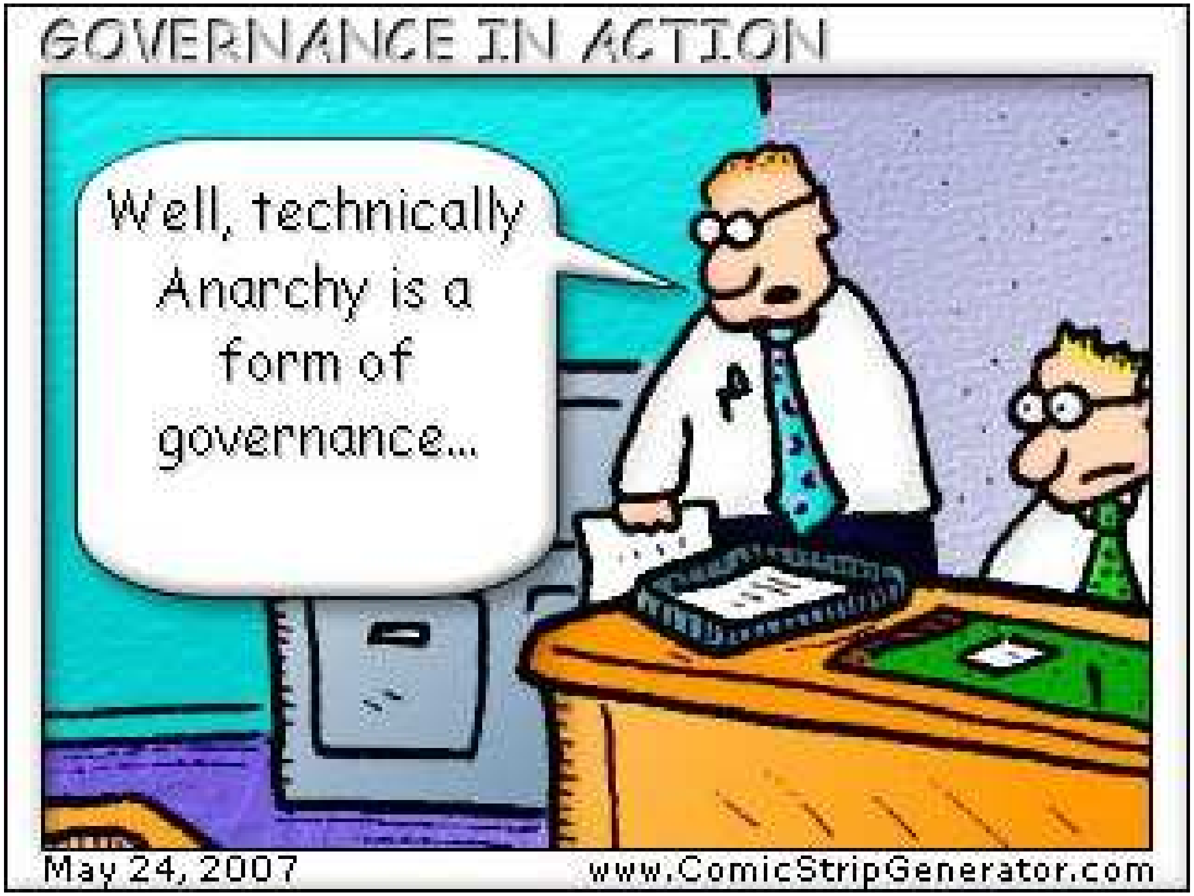
Process characterized for PROJECTS and is MANAGEABLE

Level 1 Initial

Process unpredictable, poorly controlled and REACTIVE

Source: the IBM Data Governance Council

Ha ha!


GOVERNANCE IN ACTION
Well, technically Anarchy is a form of governance...
May 24, 2007
www.ComicStripGenerator.com

Governance Maturity Levels

Level 1: Initial

Policies around regulatory and legal controls are put into place. Data considered “critical” to those policies is identified. Risk assessments may also be done around the protection of critical data.

Level 2: Managed

More data-related regulatory controls are documented and published to the whole organization. There is a more proactive approach to problem resolution with team-based approach and repeatable processes. Metadata becomes an important part of documenting critical data elements.

Level 3: Defined

Data-related policies become more unambiguous and clear and reflect the organization’s data principles. Data integration opportunities are better recognized and leveraged. Risk assessment for data integrity, quality and a single version of the truth becomes part of the organizations project methodology.

Level 4: Quantitatively Managed

The organization further defines the “value” of data for more and more data elements and sets value-based policies around those decisions. Data governance structures are enterprise-wide. Data Governance methodology is introduced during the planning stages of new projects. Enterprise data models are documented and published.

Level 5: Optimizing

Data Governance is second nature. ROI for data-related projects is consistently tracked. Innovations are encouraged. Business value of data management is recognized and cost of data management is easier to manage. Costs are reduced as processes become more automated and streamlined.

Source: IBM Data Governance Council Maturity Model (2007)

Data Risk Management Framework Maturity Levels

Level 1: Initial

There is no formal high-level risk assessment process in place. Risk assessments are done on an as-needed basis, but not yet systematically integrated into strategic planning.

Level 2: Managed

Some lines of business have processes and standards for performing risk assessments. Risk assessment criteria are defined and documented for specific items (such as credit risk) and the process is repeatable. There is limited context to validate that the risks identified are significant to the organization as a whole.

Level 3: Defined

Enterprise-wide adoption of risk assessments for specific items. Example: The privacy risk of a third-party vendor relationship uses a common scoring methodology. Risk assessment criteria are defined and documented for specific items and the process is repeatable. Data on risk assessments is aggregated for senior management. Risk assessments "outside of norm" are reviewed.

Level 4: Quantitatively Managed

Enterprise-wide adoption of high-level risk assessments for all components of the organization such as new projects, products, technologies, vendor relationships and/or applications and systems exist. Risk assessment criteria are defined and documented for all items and processes are repeatable. Data on risk assessments is aggregated.

Level 5: Optimizing

A consistent controls framework exists and is customized to the specific profile of the firm. Output of the controls assessment process is integrated into incident, reporting, and customer notification processes. A formal, ongoing high-level risk assessment process exists.

Source: IBM Data Governance Council Maturity Model (2007)

Major Components of the Data Governance Maturity Model

Category	**Description**
Organizational Structures and Awareness	Describes the level of mutual responsibility between business and IT for data governance, and recognition of the fiduciary responsibility to govern data at different levels of management.
Stewardship	Stewardship is a quality-control discipline designed to ensure custodial care of data for asset enhancement, risk mitigation, and organizational control.
Policy	Policy is the written articulation of desired organizational behavior.
Value Creation	The process by which data assets are qualified and quantified to enable the business to maximize the value created by the data assets.
Data Risk Management and Compliance	The methodology by which risks are identified, qualified, quantified, avoided, accepted, mitigated, or transferred out.
Information Security and Privacy	Describes the policies, practices, and controls used by an organization to mitigate risk and protect data assets.
Data Architecture	The architectural design of structured and unstructured data systems and applications that enable data availability and distribution to appropriate users.
Data Quality Management	The methods used to measure, improve, and certify the quality and integrity of product, test, and archival data.
Classification and Metadata	The methods and tools used to create common semantic definitions for business and IT terms, data models, types, and repositories. Metadata is information that bridges human and computer understanding.
Information Lifecycle Management	A systematic, policy-based approach to information collection, use, retention, and deletion.
Audit, Logging, and Reporting	The organizational processes for monitoring the data value, risks, and efficacy of governance.

Source: the IBM Data Governance Council

Homework Questions

Please read Kristin Wende’s paper “A Model for Data Governance – Organising Accountabilities for Data Quality Management” and answer to the following questions:

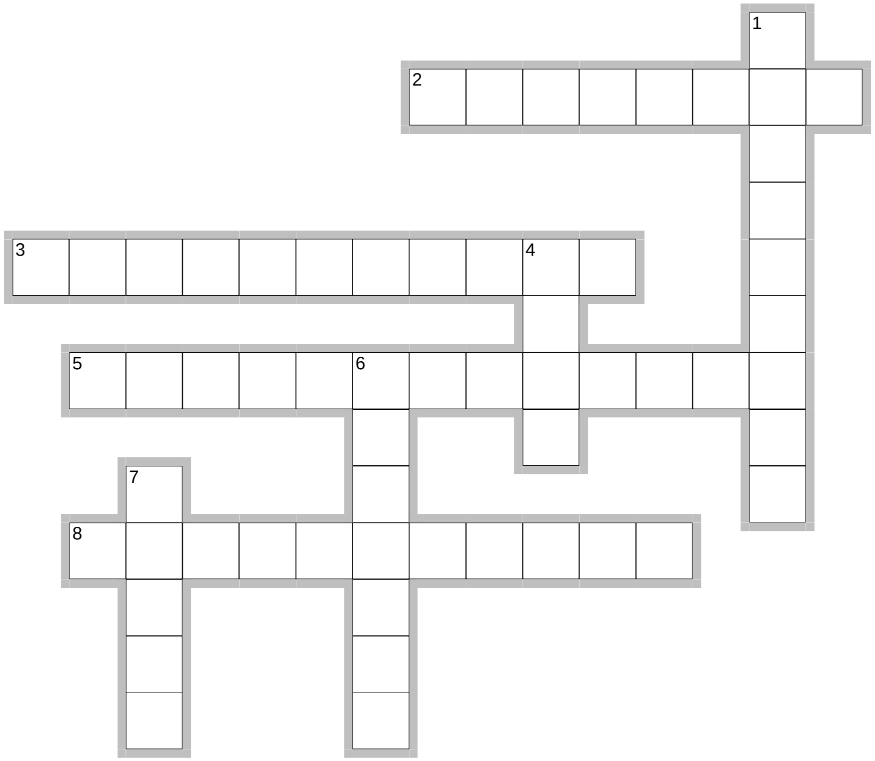
1. According to Wende, what is the prevalent *modus operandi* with regard to Data Quality Management (DQM) and Data Governance (DG) in companies today? What is the status of academic research and industry best practices? What are the particular challenges in the current approach and how does the author propose addressing them?

2. What is the difference between management and governance? How is this distinction applied to DQM and DG in the paper?

3. How would you map the data quality roles to the three DQM layers Strategy, Organisation and Information Systems?

4. How does the author relate Data Governance to IT Governance (ITG)? What ideas does she draw from ITG literature? What differences between ITG and DG does she point out?

Let’s review what we have learned today…



EclipseCrossword.com

Across

2. ________ Data Integration (CDI)

3. A quality-control discipline designed to ensure custodial care of a quality-control discipline to ensure custodial care of data for asset enhancement, risk mitigation, and organizational control.

5. MDM implementation style for reporting, analysis and central reference

8. MDM implementation style for harmonization across databases and for central reference

Down

1. System of Record and System of _______

4. Information as a service

6. SOA services are ________ coupled

7. Best practices framework for IT governance

References

- Alur, N. et al. (2009): Master Data Management: Rapid Deployment Package for MDM, IBM Redbooks
- Berson, A. & L. Dubov (2007): Master Data Management and Customer Data Integration for a Global Enterprise, McGraw-Hill
- Dreibelbis et al (2008): Enterprise Master Data Management: An SOA Approach to Managing Core Information
- Dyché, J. & E. Levy (2006): *Customer Data Integration: Reaching a Single Version of the Truth*, John Wiley & Sons
- IBM Data Governance Council Maturity Model (2007)
- Wende, Kristin (2007). "A Model for Data Governance – Organising Accountabilities for Data Quality Management", 18th Australasian Conference on Information Systems A Model for Data Governance, 5-7 Dec 2007, Toowoomba

Online References

- http://www.deloitte.com/dtt/cda/doc/content/us_consulting_im_dmrev_0606.pdf
- http://www.cio.com/article/114750/Six_Steps_to_Data_Governance_Success
- http://www.datagovernance.com/dgi_framework.pdf
- http://www.datagovernance.com/wp_how_to_use_the_dgi_data_governance_framework.pdf