

Master Data Management 101

Michael Anderson
RMOUG - February 2010

What is Master Data? What drives organizations to consider deploying a Master Data strategy? When is a Master Data management solution viable? We will explore Master Data maturity models (consolidation, co-existence, transaction, and reference). We will discuss the key business and architectural considerations. We'll do a quick review of the major PIM and CRM vendors. And we'll discuss what you can do at your Company.

MDM: What is it?

- Definitions
- Master Data Management (MDM)
 - Wiki - Master Data Management is set of disciplines, technologies, applications, policies, and procedures used to manage and govern Master Data
 - Gartner – Master Data Management is the consistent and uniform set of identifiers and extended attributes that describe the core entities of the enterprise – and are used across multiple business processes
 - Forrester – Master Data Management is a business capability enabled through the alignment of multiple information management technologies, business process improvements, and organizational commitments.
 - CDI Institute - Master Data Management is the authoritative, reliable foundation for data used across many applications and constituencies with the goal to provide a single view of the truth no matter where it lies.
 - Michael's – An application that manages the common domains of an Enterprise. The primary goal of Master Data Management is to promote common data and definitions within our organization and to reduce data inconsistency.
- Master Data
 - Wiki - Master data or reference data, is information that is key to the operation of business
- Domains
 - Enterprise Subject Area of Interest to your Organization\Business such as Customer or Product.

MDM: Why do we need it?

● Top 5 Business Drivers for MDM Solution

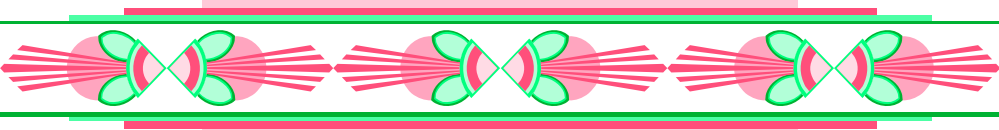
- Enable consistent usage of specific Domain related data assets across the Enterprise. It provides a uniform customer experience at the presentation layer.
- Increase operational effectiveness (i.e. Marketing - up-sell and cross-sell) via integrated DOMAIN (i.e. Customer) information
- Enable quicker solutions that drive towards a “single version of the truth”
- Mitigate regulatory and fraud risk
- Improve regulatory compliance agility

● Top 5 IT Drivers for MDM Solution

- Allows IT to further develop information management best practices and standards
- Allows IT to further develop governing support organizations to promote infrastructure and technology (Emphasis on Enterprise vs. Local Silos)
- Allows IT to further develop metadata standards (standard data definitions and data structures)
- Allows IT to create the ability to monitor and improve data quality
- Allows IT to reduce costs of integrating disparate systems of record

MDM: Business Problem Example

Hi, I would like to review my account.



OK. Do you know your account number?



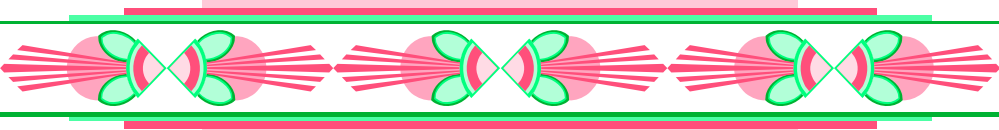
MDM: Business Problem Example

Yes, but I'm walking home and I do not have my account number memorized. My name is John Smith.



No MDM

Sorry I can't help you since we have several Customer Applications and I would spend hours trying to locate your information. Please call back later when you have your account number handy. Thank you. And have a nice day.



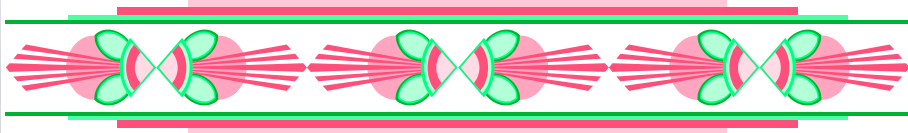
MDM

No problem. Let me look it up.

MDM: Business Problem Example

John Smith – 22 Caravel Court
John Smith – 383 Ben Avenue
J Smith - 7 East Drive
John W Smith – 99 Century Drive

I have a few John Smiths.
What's your address?



I live on **Ben Avenue**

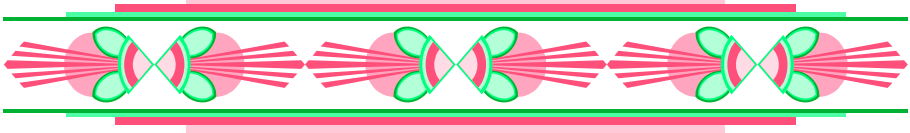


MDM: Business Problem Example

John Smith – 383 Ben Avenue
DOB: October 25, 1980
Cathy Smith – 383 Ben Avenue
DOB: January 11, 1981

What's your date of birth?

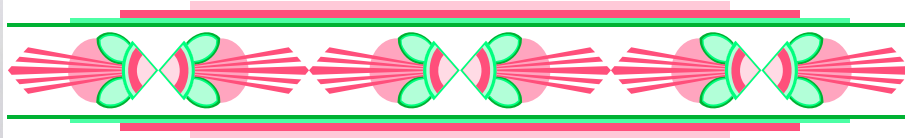
October 25th, 1980



MDM: Business Problem Example

What can I do for you?

What's my current phone bill? I've not received a bill since April.



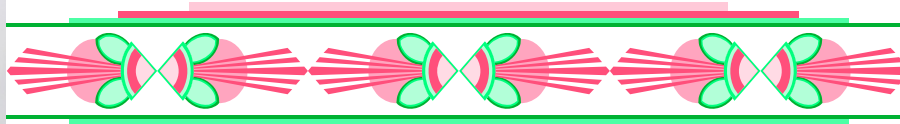
John Smith – 383 Ben Avenue
Denver Telephone Company

Products	ID	CURRENT STATUS
CELL(1)	92661	720-773-7363. \$30 due.
CELL(2)	92662	720-773-7566. \$40 due.
LANDLINE	1455	303-833-8284. \$100 credit.

MDM: Business Problem Example

You owe \$70 on your cell phones. And you have \$100 credit on your land line. By the way, we'll have a special offer on the iPhone. Would you be interested?

Absolutely! Let's talk.



John Smith – 383 Ben Avenue
Denver Telephone Company

Products	ID	CURRENT STATUS
CELL(1)	92661	720-773-7363. \$30 due.
CELL(2)	92662	720-773-7566. \$40 due.
LANDLINE	1455	303-833-8284. \$100 credit.

MDM: What's a Domain?

- Domains are subjects of interest to the entire enterprise.
- How are they discovered?
 - Engage the Business. Invoke a discovery process like Subject Area Model.
- Common Domains
 - Organizations where it can reflect any legal entity of interest, whether individual or organization
 - Examples: Agent, Customer, Employee, Prospect, Supplier, etc.
 - Product where it encompasses both physical goods as well as services
 - Examples: Part, Product, Item/SKU, Service, Product Bundle
 - Account\Contract where it includes terms and conditions, and associated relationships
 - Examples: Contract, Agreement, Transaction
 - Location where it encompasses both physical location and virtual addresses
 - Examples: Address, Geography\Region

Note. Any MDM solution must manage the Domain-related data and the relationships to other Domains

MDM: How is it different than Data Warehouse?

- Comparison : Data Warehouse to Master Data

Data Warehouse	Master Data
Primary Purpose – Store and collect data over time.	Primary Purpose – Maintain accurate and current information related to a Domain
Contains Transactional Data	Contains only Master Data
Data is historical	Data is current
Consumer of Golden Record(s). Provider of bulk data to the Enterprise Resources.	Consumer of Data. Provider of Master Data to the Enterprise Resources. Used for Reporting Purposes.
Supporting technology is oriented toward analysis and massive storage	Supporting technology is oriented toward data management
Focus on Aggregation. Broad focus.	Focus on Domain attributes. Narrow focus.
Used by Data Marts, OLAP cubes, etc.	Used by multiple operational applications and BI applications

Master Data can and may act as a *source* for a Data Warehouse.
Master Data and Data Warehouse are synergistic, and may have some redundancy.

MDM: How is it different than Data Warehouse?

- Master Data is an extension to a Data Warehouse
 - Master Data model does not typically replace the need for a Data Warehouse
 - Master Data must be fully integrated into Data Warehouse
 - Metadata integration is important
 - Data Lineage is important
 - Role of Data Warehouse
 - Historical repository
 - Provides an integration layer between disparate systems of record.
 - Provides the means for merging and reconciling systems of record.
 - Encourages\Establish pro-active data quality (monitor and repair activities)
 - Enterprise Data Source. Store and collect.
 - Provides no reporting capability
 - Role of Master Data
 - Provide current active information
 - Provide a query-able Enterprise resource. Small, quick, and adaptable.
 - Domain focused. Few key facts. (as oppose to a Data Mart)
 - Master Data will lessen demand placed on Data Warehouse or OLTP.

MDM: How is it different than OLTP?

- Comparison : OLTP to Master Data

OLTP	Master Data
Primary Purpose – Process\Collect\Manage daily transactions.	Primary Purpose – Maintain accurate information related to a Domain
Contains Transactional Data	Contains only Master Data
Data is current. Can be historical.	Data is current
Provider of Transactional Data	Consumer of Data. Provider of Master Data
Supporting technology is oriented toward transaction processing	Supporting technology is oriented toward data management
Focus on processing detailed transactional data. Narrow Application focus.	Focus on Domain attributes. Narrow Subject focus.
Consumed by multiple operational applications and Data Warehouse	Consumed by multiple operational applications and BI applications

MDM: When is it viable?

- When is a MDM solution viable?
 - Overly complex IT infrastructure
 - Application area-centric solutions are Silo-driven.
 - Slow-to-market delivery of new or enhanced solutions
 - Inconsistent definitions of corporate data assets such as Customer, Organization, Product, etc.
 - Poor data accuracy within and across business areas
 - The Business focus is on “Line of Business” data with light or nonexistent abilities to leverage information assets across lines of business
 - Redundant IT initiatives to re-solve data accuracy \ integration problems across each line of business
 - Lack of IT Budget for doing a Data Warehouse. MDM can be a quick win and less expensive to implement.

MDM: Available core architecture models?

- Master Data Architecture Maturity Models

(Based on Gartner definitions)

- Consolidation – Stand Alone Master Data

- Physically instantiated. “Golden record” single view of Domain related data stored in a central hub.
- Golden record is constructed by standardizing, cleansing, matching, linking, and merging.
- Basis for BI, but no integration back to system(s) of record
- Often considered a prerequisite for the Co-Existence and Transactional styles
- Encourages data quality improvement, completeness and semantic consistency of Master Data coming from the application portfolio.
- Batch maintenance orientation

- Co-Existence – Stand Alone Master Data and Share

- “Golden record” – same as Consolidation
- Adds harmonization of Master Data across Enterprise Application Portfolio
- Publishes selected Master Data out to the subscribing spoke systems
- Batch maintenance orientation

MDM: Available core architecture models?

- Master Data Architecture Maturity Models

(Based on Gartner definitions)

- Transaction – Integrated Master Data. Event driven.

- “Golden record” – same as Consolidation
- Transactional applications can directly read and write Master Data usually via SOA interface
- Transactional applications can subscribe to Master Data as it changes
- Real\Near time maintenance orientation for Critical Information Assets. Batch for subscription apps.

- Reference – Federated Master Data.

- “Golden record” contains globally unique identifier and very little Master Data, typically only identifying/linking attributes
- Effectively creates a new index of record that enables assembly of the “golden record” at runtime
- Assembles a single view of Master Data from various source systems in real time
- “Golden record” is typically not persisted to meet legal or regulatory requirements
- Select this model when privacy, legal, regulatory requirements prohibit data consolidation
- Requires real-time transformation of data
- Drawbacks – No focus on improving data quality and no guarantees of performance

- Hybrid Models – Imagination driven. 😊

- Budget and Scalability considerations

MDM: What are the Key Components?

- Data Management
 - How are we going to store our data? Oracle. Teradata. MS Sql Server.
What are the important characteristic? Fast query. Relative small data sets. Robustness is dependent on Domain.
How will it be organized? Third normal form? Flat like in a Dimensional Model with common set of attributes?
Must identify primary facts of interest to the ENTIRE enterprise.
- Extract and Transform data
 - Must consider software toolset to extract, transform, and load data. It may perform data cleansing.
 - Perhaps our OLTP does not cleanse address data.
 - Perhaps, the data comes from a 3rd party (Dun and Bradstreet, First Data Bank, etc.).
- Integrate structured data with unstructured content.
 - Associate unstructured contents with its Domain. Example being long product descriptions or video clips with Product images.
- Provide information to various users
 - Ability to provide Master Data to Web Sites, OLTP Application as Lookup information, etc.
 - Ability to offer end users tools like internal Search (Endeca), Workflow, and Reporting
 - Ability to provide end users with Data Lineage and other related metadata
- Synchronize with other systems
 - Ability to synchronize Master Data with both internal data sources and external destinations
 - Silo'ed processes, applications, and LOBs

MDM: What are the Key Components?

- Architecture
 - Must establish basis for single version of the truth
 - Architecture establishes integration of the toolsets and associated metadata
 - Architecture and Business determine source of record(s)
 - Without integration being considered, there will be multiple versions of the truth
 - Toolsets should be integrated through metadata
 - Must establish basis for Data Quality and Data Auditability
 - Architecture and Business must establish how data quality will be measured
 - Must determine approach for handling bad data. And process for making it good.
 - Eliminate direct manual intervention. Keep SAS70 in mind.
 - Must determine approach for data lineage and auditability
 - May want to consider Data Stewardship mechanism
 - Must establish appropriate architectural layers
 - MDM architecture should be composed of multiple layers
 - Determine purpose of each architecture layer
 - Rules of engagement for each layer
 - Performance expectations for each layer
 - Architecture must foster\define development standards and design patterns
 - Platform requirements
 - Business requirements determine hardware architecture
 - Architecture defines platform requirements
 - Requirements are traceable
 - Metadata Architecture
 - Centralize metadata strategy integrates all the individual components of the MDM application
 - Lack of consideration of Metadata Strategy is a major cause of MDM failure
 - Metadata : Central vs Local
 - Project Implementation Approach\Strategies
 - Data Governance Approach\Strategies

MDM: What are the Key Components?

- **Business**

- Discuss requirements for derived Master Data
- Discuss requirements for out of order (late arriving) fact related data.
- Discuss requirements for SLA (Service Level Agreements) and RTO (Recovery Time Objectives)
- Discuss how effort will be resourced: RPO (Partial or full outsourcing)
- Discuss data integration requirements
- Discuss data cleansing requirements
- Discuss data synchronization requirements
- Discuss Source of Record\Golden records
- Create SAM for the Enterprise
- Use a discovery process
 - Create\Revise Subject Area Models
 - Identify key business requirements (wide not deep)
 - Assess business process maturity level

- **Development**

- Determine SDLC: Agile vs. Waterfall methodology
- Establish development strategies and design patterns prior to deploying any methodology
- Commentary: If you chose an Agile approach – Iterations (sprints) should be dedicated to creating business functionality, not dealing with architectural or development issues

MDM: Keys to Success

- High Quality Information
 - Accurate.
 - Auditable. Trusted.
- Project Management
 - Focused scope. Narrow. Select your DOMAIN.
 - Identify important business requirements
 - Initial investment in proportion to value delivered
 - Strategic to Tactical. Planning is the key.
 - Plan of action designed to achieve a particular goal
 - Detailed actions to achieve objectives determined by strategy
- Business and IT strategic alignment
- Management Support
- Teamwork

MDM: Efforts at your Company

- Corporate Game Plan
 - Business Directives
 - Is there a compelling Business Reason(s) to consider Master Data?
 - Are there sufficient resources (Time and Money) within your Organization to undertake such an effort?
 - Sufficient skill sets? Data Modeler? ETL experts\Developers? Enterprise SME? Project Management?
 - Is there a sufficient Enterprise commitment to undertake effort? Who's your Business Stakeholder\Champion?
 - Create Business Case. Emphasize ROI, TCO, Cost Savings, etc.

MDM: Efforts at your Company

- Corporate Game Plan
 - IT Directives (Part 1)
 - Create SAM with an Enterprise orientation
 - Excellent resource:
 - Steve Hoberman's Data Modeling for the Business: A Handbook for Aligning the Business with IT using High-Level Data Models
 - **Review the SAM's definitions and relationships with all parties to ensure agreement (Is a Customer a Customer)**
 - Identify the COMMON Enterprise Domains

MDM: Efforts at your Company

- Corporate Game Plan
 - IT Directives (Part 2)
 - Determine potential Master Data candidates
 - Select a primary candidate based on
 - Best ROI
 - Infrastructure – Alleviate COMMON request or repetitive queries
 - Improve Customer Service\Experience
 - Increase Revenue opportunities
 - For the primary Master Data candidates, determine COMMON DOMAIN attributes and major facts
 - Discuss primary Master Data candidates with the Business to determine best candidate
 - Discuss Source of Record\golden records for the candidate
 - Identify potential Source of Records
 - Identify “Source of record” owners
 - What is it? Oracle, Flat Extract, MS Sql Server, etc.
 - When is available?
 - Etc.

MDM: Efforts at your Company

SAM: To be provided – AS IS

SAM: To be provided – TO BE

Review Application Subject Area Models – Identify potential DOMAIN candidates - TBD

MDM: Who are the MDM Vendors?

- MDM Vendors
 - CRM(Customer Relationship Management)
 - Ardexus MODE v. 5.5 -- Ardexus Inc.
 - C2 CRM v. 8.0 -- Clear Technologies Inc.
 - CMS/Oncontact V, v. 5.2 -- Oncontact Software Corp.
 - Data Flux
 - Goldmine v. 7.0.3 & HEAT -- FrontRange Solutions Inc.
 - IBM WebSphere Customer Center – IBM
 - Initiate Systems
 - Maximizer Enterprise 9.5 -- Maximizer Software Inc.
 - Microsoft CRM 3.0 -- Microsoft Corp.
 - NetSuite CRM v. 11.0 & NetSuite v. 11.0 -- NetSuite Inc.
 - Siebel UCM - Oracle
 - Powertrak v. 8.02 -- Axonom Inc.
 - Purisma
 - Relavis CRM 7.0 -- Relavis Corp.
 - RightNow CRM v. 8.0 -- RightNow Technologies Inc.
 - Sage CRM 100/200 v. 6.0 -- Sage Software
 - Sage SalesLogix v. 7.0 -- Sage Software
 - Salesforce.com -- Salesforce.com Inc.
 - Salesplace 2006.3.2 -- Interchange Solutions
 - SalesPage CRM -- SalesPage Technologies LLC
 - SAP
 - Siebel CRM OnDemand -- Oracle/Siebel Systems Inc.
 - Siperian
 - StayinFront CRM v. 9.3 -- StayinFront Inc.
 - VisionWare

MDM: Who are the MDM Vendors?

- MDM Vendors
 - PIM (Product Information Management)
 - Amphere Solutions
 - Comergent Technologies
 - Flow Systems
 - FullTilt Solutions
 - GXS
 - Heiler Software
 - Hybris
 - IBM
 - Inovis
 - I2 Technologies
 - Riversand Technologies
 - Stibo Catalog
 - Velosel

MDM: Summary

- Summary
 - Master Data provides an integration point for Chaos
 - Master Data is focused on an Enterprise Domain
 - Master Data stores few key facts and attributes of interest to all Lines of Business for an Enterprise
 - Master Data encourages Enterprises to self reflect (Who we are?)
 - Master Data is focused on NOW (current active records)
 - Master Data Management solutions can be replicated
 - Master Data Management solutions can be purchased
- Questions and Answers