International Journal of Scientific Research and Management (IJSRM)
||Volume||09||Issue||12||Pages||EC-2021-715-729||2021||
Website: www.ijsrm.net ISSN (e): 2321-3418
DOI: 10.18535/ijsrm/v9i12.ec06

Architectural Frameworks for Master Data Management (MDM): Enabling Holistic Customer and Product Master Synchronization in Investment Portfolios

Ramesh Inala

Data Engineer, Prudential Financial, Newark, NJ,
ORCID ID: 0009-0009-2933-4411

Abstract

Master Data Management (MDM) is a critical discipline for enterprise architecture. MDM is ubiquitous and multi-sourced. It spans and integrates systems and data domains, working closely with Business Process Management (BPM). It is the prerequisite of Business Intelligence (BI), and of business derived from the dynamics of the digital universe and from rich Data Lakes. It asks, and answers, unique questions about the business. For a robust MDM implementation, the discipline needs to be backed with an overarching MDM framework that engages business and IT. Such an MDM architectural framework is incomplete without the architecture of systems connected to master data, and the pivotal role of these systems in the business process. We explore ideas and insights behind the MDM architectural frameworks, and related questions the frameworks answer.

Master data is a specialized class of data of any business entity for an organization, the context in which the business entity exists, the why pertaining to the business entity. Within a business, it is core and is at the heart of the business process architecture. Data integration and consistency is a primary goal of data management. The objective of MDM is to establish a single version of the truth of integrated master data that is made available in a consistent manner to multiple consumers, internal or external to the organization. MDM is the act of creating and maintaining a master brownstone. MDM requests this silos view and develops it further with a multi-silo and multi-replication perspective to build the integration framework for a given enterprise. MDM is the responsibility of the business people owning the process, but working with the IT organization or a third-party vendor to implement the physical data structure and make it available to the users.

Keywords : Master Data Management (MDM),Architectural Frameworks, Data Governance, Customer Data Integration, Product Master Synchronization, Investment Portfolio Management, Data Quality, Reference Data Architecture, Holistic Data View, Metadata Management, Data Lineage, Data Consolidation, Entity Resolution, Information Architecture, Financial Data Standardization.

1. Introduction

The collection and analysis of data is a key element to the viability of contemporary organizations. Organizations generate huge amounts of operational data on a daily basis, which merely serves its initial purpose of facilitating ongoing discrete transactions. However, as organizations attempt to leverage these data assets to obtain useful information which reflects the entire organization and which can be used to decisively impact various strategic issues, concerns arise about the quality and effectiveness of the data. Organizations face a myriad of problems with the operational data. The data are typically kept in numerous independent systems, contained in distributed databases which are often heterogeneous. The data can range in age from just generated to many years old. The data may also be incomplete, inaccurate, insufficient, inconsistent, redundant, and inaccessible. Furthermore, organizational data, such as customer, product and supply chain data, are increasingly more complex than earlier databases allowed for since they are typically hierarchical and interrelated using multiple types of relationships.

There is primarily interdependence but little integration between data used by departments for specific departmental applications. Interfaces to integrate the various legacy systems are often complicated and inflexible. Furthermore employees who access these systems must learn numerous user interfaces in order to utilize different systems. However, despite these challenges, the potential for improved company performance is great for organizations that deduce useful information from their operational data to make organization-wide strategic decisions. These organizations should be able to tailor their products and services to better meet the specific needs of their customers, thereby reducing costs and boosting profitability.

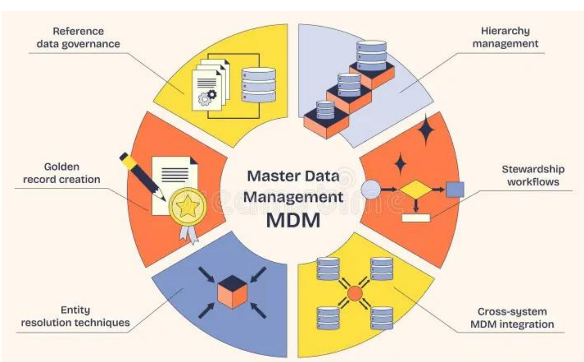

Fig 1: Master Data Management

1.1.Background And Significance

Master Data Management (MDM) refers to the processes, governance and tools employed to ensure a consistent, accurate and authoritative view of an organization's critical business data. By maintaining a trusted source of reference data in the form of a data repository or hub, MDM enables reporting, querying, analytics and business intelligence applications with good quality data. As organizations focus increasingly on improving the quality of their business information, MDM has emerged as a critical component of enterprise architecture. The failure to efficiently manage business-critical data over its life cycle will, sooner or later, impact the credibility of an organization. The core concept of MDM is enterprise scale data modeling of reference data such as accounts, customers, products, employees, locations, channel partners and transactions. However, enterprise data model design is just one aspect of addressing a number of business and technical issues to deliver an MDM solution. It is an issue with the potential for far-reaching consequences across multiple domains, underscoring the value of an organization's investment in a clearly defined and supported enterprise architecture framework. This paper describes a four-layered, industry-agnostic architectural framework for MDM. The framework also leverages best practices. We expect to organize a more detailed MDM implementation guide and user experience guidelines for each of these layers later.

2. Importance of MDM in Investment Portfolios

Investment management drives value creation for customers. It is, in fact, an organization's means for exchanging value with its customers. Obviously, asset allocation and portfolio construction are core functions. Risk diversification, asset selection, and investment strategy are part of the investment decision-making process, nicely supported by research and analytics. Bringing these customer-facing activities to life are the equally critical operational functions, which provide execution and control support. Transactions, such as trading, settlement, revenue and cost management, accounting, and reporting, need to be performed with speed, precision, and risk control. Everyday purchases must balance return on investment against operational costs. To do away with redundant investments and duplicate costs, all this performance activity must pull on a common data resource. This collective body of information, modeled sensibly, and reconciled regularly, is referred to as master data.

Investment products and clients comprise the master data pool. Portfolios are the focal point around which all investment activity revolves. So portfolio-related data is critical. Ninety percent of buyer behavior is governed by emotions, primarily, in the form of perception-driven motivation, and investment decision-making is no different. For Money Managers, a bad perception will lead to poor sentiment, which, in turn, will translate into an avoidance mode. A Media Council that has a bias against the Money Manager will chisel at its trustworthiness, credibility, and integrity, leading to a credibility crisis. Concomitant with this crisis will raise the hue and cry of indignant news coverage. Conversely, the Money Managers that excel at risk maximization will thrive. These are the Money Managers that have correctly identified the investment needs, sentiments, and priorities of clients and done their utmost to exceed the latter's return expectations and risk tolerance levels.

Equ 1 : Unified Master Equation

$$U = f(CM, PM, GR, DS)$$

Where:

- **U** = Unified master view across the portfolio
- **CM** = Customer Master data layer
- **PM** = Product Master data layer
- **GR** = Governance & Roles
- **DS** = Data Standards & Schema mapping

3. Key Components of MDM Frameworks

Architectural frameworks often expose the necessary aspects of their properties and highlight their relationships. A reference MDM framework is not prescriptive but allows for an example to discuss and explore the options available.

A conceptual version of a reference MDM framework with three key domains, called data governance, data quality management, and data integration, represents the topical specialization of teams that participate in an MDM initiative. They either realize MDM activities independently or realize it together using shared interaction points. Moreover, each domain has specific requirements for the generic principles of Enterprise Architecture found in realization projects.

Within this reference MDM framework, data governance manages the arbitrary complexity of enterprise semantic models using a surprisingly small number of support items. This complexity is arbitrary because creating a universal, fixed semantic model is not feasible. Enterprises need a tailored semantic model that differs from all other enterprises in some areas. Only the most specialized items are managed by the governance domain, called enterprise Master Data Items, which cross domain and enterprise boundaries and are shared by successions of business processes used for core business. Other items that form business process specific structures and cross a small number of domain boundaries are called domain MDIs and are covered by Data Domain Specialists. These are managed by Data Domain Management. All other items important to certain business processes but presenting no risk of requiring long-term relationships forming core business are not managed at all but covered by Data Executive Control.

Fig 3: Components of MDM Frameworks

3.1. Data Governance

Data Governance refers to the people and processes that provide for intelligent and effective use of data across an enterprise. An organization’s data

governance program defines how data is consistently represented so that it can be used effectively, while minimizing risks associated with its use. It also lays out the policies and procedures that ensure the right controls are in place to assure that data is trustworthy. Throughout this document, we will refer to a "data governance program." This term covers both the technology and the policies and procedures that drive the business processes. Data governance provides the tools and processes to manage enterprise data assets in a compliant manner. These tools simplify the execution of the data governance policies and the data governance procedures documented using data governance. These are not meant to be done in one step, nor considered complete and ignore the best practices that involve continual status assessments and adjustments of the organization's data governance program. These steps are done regularly to ensure that the data governance program continues to align with the organization's goals and objectives, while also being able to react to changes in the organization's strategy and direction.
A data governance program has processes for establishing and managing data stewardship. A data stewardship process is made up of three interrelated functions: assignment of data stewardship roles and responsibilities, the routine activities done by business data stewards to fulfill their roles and support enterprise data as an asset, and the communication and approval workflow supporting stewardship decisions. Each of those stewardship functions plays a critical role in how effectively an organization can manage data as an asset. Formalizing the assignment of roles and responsibilities ensures that business data stewards know what data they are responsible for.

3.2. Data Quality Management

Overall, there is little argument that ensuring the quality of master data is a key driver for a successful Master Data Management solution. Data Quality Management provisions the creation, and real-time monitoring of data quality for critical master entities, including Customers, Product, Location, and Supplier. In addition, it is common that additional "extended" entities are included under the DQM umbrella such as Party, Contract, and Structural. DQM enables consistent policies and processes for managing data quality, clarifying which party is involved in defining data quality strategies and overseeing management.
Data Quality Management allows organizations to define the principles that establish the rules defining first how data needs to look and second how it can get there. Data quality provides the framework containing the following components, either explicitly or implicitly concerning (1) Processes - the practices or activities followed to build, validate, and deploy master data; (2) Quality Dimensions - the characteristics against which data is measured for quality; and (3) Quality Metrics - the quality dimensional measures for MDM that help to assess organizational compliance. Quality dimensions include data accuracy, completeness, consistency, efficiency, integrity, precision, reliability, relevance, simplicity, timeliness, and understandability.
In addition to the foundational framework of DQM, it is expected that additional functionality is needed that provides the definition of how the data needed to look and more importantly how the data would get cleaned, uploaded, and help with ongoing maintenance once the data was in the MDM repository. Because master data is central to the overall strategy of MDM, and needs to be of very high quality, organizations have developed very complex processes and systems for Data Quality Management. Much of this complexity is due to the strong dependency on many different systems, applications, and processes. Each organization has to manage their unique set of sources for master data, as well as determine the responsible parties for onboarding, maintaining, and collaborating for data and process quality throughout the enterprise.

3.3. Data Integration

MDM represents a specialized function of enterprise integration services. However, MDM consolidation based upon master + detail relationships distends enterprise data integration, extending its temporal dimension. At the same time, this information lock is a fundamental practice in data stewardship, governance, and architecture, as it relates to its regulation, informing actions upon competitive advantage derived from creating customer knowledge and efficiency based upon B2B relationships. Automated execution of the data access, preparation, staging, and transmission episodes involved in the integration of source data into the data repository, or ‘hub’, is not achievable through traditional data warehouse extraction, transformation, and load services, which reflect batch processes responding to defined delivery schedules. The data in MDM repositories – customers, products, assets, locations – may change at any time. Synchronization of all actors in an MDM ecosystem around the common definition of a customer, product, or warranty document is crucial because many organizational functions leverage it. Master data-enabled enterprise functions include all areas of finance, sales and marketing, product, production, customer service, and maintenance. Enabling on-line and offline transactional and analytical services, point-of-sale and periodic data collection acts by organizations, suppliers, and customers continuously reflect external change upon internal environments and vice versa. Standardized, specifically defined, and regulated interfaces support the replication, enrichment, and dissemination of this dynamic information flow involving detailed reporting and high business process throughput.

The integration approach to data presents the alternative of master data expertise and synchronizing details into repository content, or a point-to-point architecture, where each domain transaction system is directly managed by the enclave network architecture and rules governing its respective source data management. Either alternative requires the data sources be sufficiently authoritative, controlled, and adequate as to guarantee the integrity of the repository content.

4. Architectural Models for MDM

In the context of MDM, the Architectural Models describe its various implementation options and techniques based on the spatial distribution of MDM capabilities, data, and processes. These models act as a foundation for an MDM solution. They define how the enterprise approach towards its information management, data governance strategy, and user buy-in to MDM Organization, Business Processes, and Enterprise Information categories are modeled and used. Thus they could also be termed "architectural models for an enterprise MDM solution and data governance." Based on MDM capabilities, data, and processes, there are three MDM architectural models: Centralized Architecture, Decentralized Architecture, and Hybrid Architecture. An organization can choose any or a combination of these architectural models to meet its business requirements.

The centralized MDM model has a dedicated MDM system and repository holding the golden copy and performing most of the MDM functions, such as data synchronization and distribution to backend stakeholders. Data in the other enterprise project systems are often considered shadow copies. While the MDM hub may or may not be tightly integrated with enterprise project systems, the majority of MDM functional capabilities can only be performed through the MDM hub. Data synchronization is often one-way, because the MDM hub often enforces strong data governance policies on other systems and has a higher data quality compared to them. The tapered ice-cream model describes this architecture well. However, Enterprise Project Systems that hold the shadow copies can still cling tightly to the MDM hub.

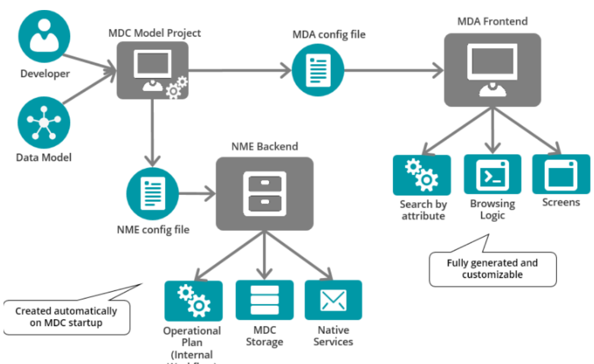

Fig 4: MDM Model

4.1. Centralized Architecture

To ensure operational efficiency, many companies have traditionally relied on centralized computer architecture. This model is logically an extension for the MDM solution. Data is extracted from the source systems, cleansed, transformed, and consolidated into a single database that is a trusted source for authoritative reference information. Business intelligence solutions extract the data from the MDM hub and deliver it to the various transaction processing source systems used throughout the organization.

The centralized model is attractive for an organization because it provides a single source of master data that can create a "single version of the truth." Having a single validated version of the master records stored in the MDM hub provides the best opportunity for the proper integration of the extensive business transactions generated throughout the organization. The integration process has severe consequences when duplicate records exist for the same business entities located in separate source systems and integration solutions. A company that retains its customer master records in printed telephone directories and utilizes them as the primary data source for accessing customers has been a long-time industry curiosity. Clearly, duplicate telephone entries for the same individual generates business transaction problems for the organization.

Other advantages of centralization include simplified regulatory compliance and audit trail reporting, and the standardization of the process of collecting and storing master data. A single MDM product configuration with uniform metadata and management support functions is configured and applied to all critical business processes that access and share master data. A standardized product configuration creates a business environment that minimizes the specialized skills required for the teams that administer the different source systems using the same MDM system. An additional factor is the fiscal cost to the organization for operating a mastery system that is decentralized. Centralizing the MDM function results in fewer licenses to acquire and ongoing maintenance fees with the software vendors.

4.2. Decentralized Architecture

The knowledge of how information is used within the enterprise provides the necessary insights to define an adequate and reliable MDM solution. The information used across the enterprise often is not known and accounted for in a central schema. This normally is the case for reference data, since it is affected by a number of different criteria, like differing stakeholder requirements, regulation differences, cultural differences or localization constraints. These decentralized and localized differing requirements are best served by enabling domains to autonomously manage their own reference data. However, for reference data to serve its purpose, a minimum of control, accountability, validation or certification is required to secure enterprise interoperability and performance. A decentralized MDM solution needs to enforce this accountability in a robust enterprise governance model for optimizing both impact areas. Bear in mind that while MDM may be realized in a decentralized manner, it needs to implement interfaces to expose this information to other domains and allow for the MDM functions or services defined in the enterprise MDM governance model. There are a number of decentralized MDM functions and services offered to domains that govern local MDM implementation. They include domain reference data management, business activity monitoring, enterprise process management, business intelligence integration, domain master data establishment, domain master data maintenance, domain master data federation, validation and certification. With respect to service granularity, these services may be implemented at different levels.

From a service perspective, a decentralized approach options from domain-centric federated MDM, domain-centric replicated MDM, domain-centric federation and validation MDM, domain-centric retraction MDM, domain-centric shared MDM, domain-centric certified MDM and domain-centric external MDM services. All of these options need to be well accounted for when developing the enterprise governance model. However, the enterprise governance model defining how the services listed above are being executed is largely work product dependent, especially when it comes to identifying which organization is providing these services. The domain-centric MDM model aims to reduce costs associated with the storage and management of shared MDM work products by defining a much smaller number of shares while at the same time enforcing high levels of response time to the consumers.

Equ 2 : Data Quality Score Equation

$$DQS = (AC + CM + CC + TC) / 4$$

Where:

- **DQS** = Data Quality Score
- **AC** = Accuracy
- **CM** = Completeness
- **CC** = Consistency
- **TC** = Timeliness

4.3. Hybrid Architecture

The hybrid architecture allows for a combination of centralized control with decentralized processing. Examples might be a worldwide conversion done at headquarters with appropriate systems transferred to corporate offices for local data collection via data entry terminals, or a master data list maintained at headquarters but sent out running on local facilities which update the list via telecommunications. The hybrid, which has inconsistency because coordination for those portions that are centralized is missing, shows promise of broadening the applicability and feasibility of MDM relationships and extends the family of suggested systems. Design choices often come down to either total concentration, which can have unfortunate time-sharing results on performance, or local solutions to autonomy problems, which lead to today's chaos. Therefore, there is an increasing interest in multi-user systems whereby a central file can reside on a central processor with users operating independently when demand is light but ability to rapidly connect for central control of performance bottlenecks and online editing. Because much of the research on multi-user systems has been analytical, it is relevant that there appear to be more central processors being built capable of operating in a time-shared, multi-user manner. However, the added power has often not been leveraged sufficiently to support transaction processing rather than only desktop computing. The limitations of present time sharing systems to handle high-volume transaction processing is one of the strong arguments for hybrid as opposed to total centralized processing of supporting information.

5. Technologies Supporting MDM

Both current and future MDM initiatives can benefit from various software technologies that increase the quality and lower the cost of MDM solutions. MDM and related initiatives such as data warehousing, data virtualization, data security, customer experience management, and regulatory compliance are usually implemented in the context of enterprise architectures designed to support core business processes. The MDM architecture leverages existing and new data sources and repositories to deliver authoritative master data to applied BI or data integrity solutions. The MDM architecture is focused on ensuring that the right data is delivered at the right time to authoritative source systems, reference applications, and data consumers to execute and optimize enterprise processes.

MDM repositories can be implemented using traditional RDBMS solutions as well as specialized DBMS systems. Each of these DBMS options has its strengths and weaknesses when it comes to the specialized requirements of MDM implementations. General-purpose RDBMS solutions tend to deliver very high data availability, load performance, and transaction support for operational master data repositories which implement transaction workload MDM use cases. At the same time, specialized MDM database engines tend to dominate in performance and scalability for data warehouse and applied BI refresh cycle MDM use cases.

ETL has been an integral technology in many past MDM implementations and will likely be an even more prevalent component of future MDM scenarios. Currently deployed, mature, and optimized for performance, reliability, and precision, ETL solutions are almost ubiquitous in MDM architectures. The development of enterprise applications for building transformation mappings in today's competing ETL solutions has become relatively standardized. However, ETL vendor approaches differ in many other areas including overall architecture, scheduling, monitoring and management, pricing, built-in data quality capabilities, and support for data de-duplication and matching.

Fig 5: Mobile Device Management - Mobile-Technologies

5.1. Database Management Systems

Master Data Management (MDM) is a business-driven discipline in charge of the creation and maintenance of a single source of truth for the entire organization. IT then provides the technology and infrastructure as support to the business initiatives. A considerable part of MDM is to ensure that the master data is made available in a trusted format for operations, analytics, and sharing for any of the stakeholders in the data.

A traditional database, such as a relational database or a NoSQL store, is where the business and technical needs meet and combine in order to create a high-level implementation of the business objectives. In the MDM domain, we can find Polar Lake, a solution developed by a Canadian company which focuses more on data preparation than on being a typical central repository of master information. It can be used for loading master data in some of the typical DBMS and Data Warehousing solutions available in the market and may fulfill the central repository requirement for certain customers in niche markets.

Niche vendors also offer many solutions in the MDM space but they have chosen not to include a data repository as part of their offerings, effectively providing a data management solution, and have entrusted the data repository layer to other vendors. An example of this type of solution is the addressed implementation in which a data repository is utilized. They combined the data management capabilities of the data integrator and the integration capabilities of the business intelligence tools. A single MDM logical model was developed that made it possible to clean, integrate and harmonize the address data.

5.2. ETL Tools

ETL tools are the main technologies for building MDM system data repositories and loading data into those repositories. By extracting data from transactional systems or staging areas, transforming the data to align with MDM system data model design and physically loading the data into the MDM system, ETL tools support the initial and ongoing population of the data repositories. Some refer to ETL as Extract/Transform/Load or Load.

The definition of an ETL tool includes the ability to extract, cleanse, standardize, enrich, match, and merge. While MDM is increasingly becoming a part of the general DM conversation, these emphasize its special MDM aspects. Considering the ETL function only, the data in the MDM repository is to be loaded from some source, most likely an OLTP database. It is often asserted that the source data is dirty because it contains many errors, which is the cause of the MDM solution. However, it is incorrect to consider that the only goal of ETL processors is data loading. Sometimes ETL also has to re-expose the data from the MDM system into some OLTP for the reason of reuse in their solutions. In those cases, both loading and publishing should be taken into account.

As such, while data loading is the most important function of an ETL tool from an MDM perspective, ETL tools have additional functions related to both aspects, especially in terms of data cleansing and enrichment. ETL tools allow an organization to profit from reusing the cleansed and standardized data with errors and enrichment that would also not be uploaded into the MDM system.

Equ 3 : Governance Efficiency Ratio

GER = E / (R × G)

Where:

- **GER** = Governance Efficiency Ratio
- **E** = Number of resolved data conflicts
- **R** = Number of rules applied
- **G** = Governance overhead (person-hours or system cycles)

5.3. Data Virtualization

Data Virtualization is a recent technological development that further blurs the line between ETL tools and Database Management Systems, making it sometimes difficult to identify what category of MDM support tools they identify with. We include Data Virtualization tools in our discussion of MDM support tools because they have the potential of providing MDM Centralized Access Aggregation and Control Services for federated repositories of domains' data without requiring copies of the data physically stored in the MDM database or the data residing in specific data sources relational or not. Along the lines suggested by the MDM Architectural Framework, Data Virtualization tools can potentially create an MDM-like layer on top of a digital enterprise's data working as if it was its MDM database, while avoiding what may be an overloaded copy.

Data Virtualization tools can access data stored in multiple data sources, which may be relational or not, structured or not, in a coherent way that is simultaneously transparent to applications and external parties using Data Access Methods and presenting the data in a standard-dimensional tabular way. A Data Virtualization Server acts as a middleware that connects to the data sources and serves requests for data from applications and external parties through the Data Access Methods, which may use proprietary data formats or definitions of the information to describe the data. The functionality of Data Virtualization Servers usually also includes caching capabilities that may tune performance, Data Source Connection Pooling capabilities that may handle system resources' utilization, and Monitoring and Administration capabilities that may assist Data Virtualization management and administration. There are two RAD capabilities that make Data Virtualization easy to use: data source plug-ins, and Data Virtualization Projects generation based on a low-code approach.

6. Challenges in Implementing MDM

MDM is not a panacea. Organizations should realistically evaluate the complexities and pitfalls to produce a trusted golden record. MD attempts to overcome the limitations of existing techniques and technologies, however, there are inherent problems that need to be considered and addressed in implementing MDM solutions. When a business tries to make MDMa reality, they encounter issues regarding their time, people, costs, data performance, delivery, quality, and trust. Organizations undertaking a master data initiative must be aware of the potential challenges to help prioritize their resources, effort, and focus.

Organizations often encounter problems when trying to implement MDM. They discover too late that their enterprise is riddled with data silos and multiple discrete systems that hold apparently duplicate data that has been generated, modified, and created in isolation. Providing a set of standard naming conventions and standards, reducing excess copies of how data is structured, creating a definitive and unchanging data model, minimizing dependencies on systems that generate duplicate data are crucial tasks that organizations must address when attempting to implement MDM, otherwise organizations are faced with data silos that can not easily be merged.

Implementing a master data system is one of the most difficult challenges in data management. For organizations, integrating data into a unified whole is not easy. Thousands of transactional systems may collect, maintain, and utilize customer or product data. Petabytes of data may meld from these disparate units, and the tasks required for the centralization of critical components may be overwhelming. The diverse management of unique 'islands' of customer data creates massive problems for professionals and analysts, particularly in the areas of redundancy, inconsistency, discrepancy, compilation, duplication, and maintenance of integrity. In the context of MDM, the change process is quite complicated. MDM related changes may affect all data applications of an organization, and any of them may be negatively influenced by the changes. History has shown that notification is not sufficient in managing change. Due to their daily routines, employees are not primarily sensitive to impending changes that may jeopardize data quality.

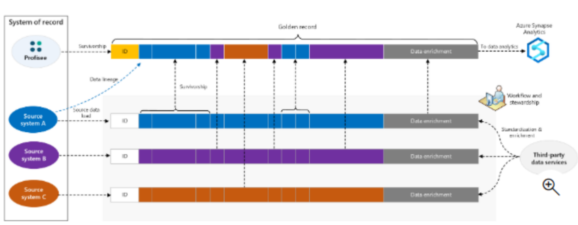

Fig 6: architectural frameworks for master data management (mdm)

6.1. Data Silos

The concept of data silos refers to the storage of data in one location or database. Data is isolated from other parts of the organization and is difficult to access and manage. For most companies, this typically happens due to organic growth, expansion through mergers and acquisitions, or the implementation of various systems to maintain different streams of business. Information or systems associated with a particular business unit, department, or function is not integrated across the organization as a whole.

Oftentimes, enterprise applications that manage the core operations of a company, such as Enterprise Resource Planning, Customer Relationship Management, and Supply Chain Management, are implemented with specific business drivers in mind. These applications are built to support the specific needs of the department, but they do not integrate with other internal systems, such as point-of-sale, telephone support, or accounting systems. Over time, companies with particular care taken to improve efficiency and visibility into business operations through enterprise applications still end up with fragmented data that is duplicated across many systems and cannot deliver enterprise-wide visibility. And while enterprise applications have enabled organizations to standardize processes and improve operational efficiencies, these applications have usually not enabled data integration across the enterprise, let alone integration with external systems, partners, or customers.

Today, consumer and partner data exists in a variety of functional systems, such as ecommerce, sales, marketing, reservations, and fulfillment. No one functional capability can capture the entire person or customer experience from initial interest through post-sale fulfillment. Each interaction delivers part of what we need to know about that individual, but those interactions typically reside in separate data stores, further capturing the data within departmental silos. The data in these silos needs to be related to deliver a complete view of the individual and their interaction with the organization.

6.2. Scalability Issues

Standard enterprise MDM implementations first seek to reduce the number of data identification keys to a manageable figure. MDM then attempts to aggregate all data sharing the same data identification keys while enforcing that the merged records link all its keys' associated sources. Next comes the issue of integrating the MDM implementation into the everyday information exchange processes to keep MDM data fresh and up to date. But once these steps have been taken, the next challenge comes along: allowing MDM to deal gracefully with the next source whose data volume suddenly increases dramatically. Such happenings, like year-end holiday shopping or tax season followed by refunds, reduce but are not limited to the retail, insurance, and taxation industries, respectively. If MDM fails gracefully, these sudden volume increases will necessitate service interruption, which is especially critical if MDM data supports constant operations in other industries and sectors that do not sit idle during phased seasonal business cycles.

In approving and finalizing an MDM vendor's products and services, operational resiliency aspects must be addressed. Standard product offerings must demonstrate cost-effective data building block scalability accordingly. Staff size should initially reflect system design with a number of concurrent connections from source systems of the same size. Business process and design requirements must include the monitoring of implementation during the product ramp-up. Data migration backup processes must reflect the estimated minimum duration for this ramp-up stage. Data translation and transformation support for diverse data structures from different source system providers at differing and changing update intervals must demonstrate scalable speed. Would-be MDM customers must remember too late what they were sorely lacking if this capability is not proven in support of customers. Otherwise, they might pay large supplementary-use support charges, pay high training or service costs to facility in-house staff, or cope with ongoing data quality problems and poor electronic trading partner relations, for years after going live with MDM. The business impact is substantial too.

6.3. Change Management

Master Data represents the vocabulary in which business communicates. The business functions change over time which affects how the people within the business want and use Master Data. Making a change to a Master Data Attribute directly impacts how the business interacts with applications that manage transactional data. The authoritative Model is periodically updated to ensure that it reflects current requirements of the business stakeholders. These requirements state the elements that should be represented, the systems in which they will be used, and the business rules that manage their use. The business rules will also determine the frequency of updates in addition to the selection of values and the application of those values to records used by transactions being managed. It should be apparent from the above that a Master Data problem can occur quite regularly without any IT interaction. The above serves as an example of an effect of change management.

Managers from all the business functions must interact with the Model, on a regular basis, to ensure that the attributes remain relevant. Every time a change occurs, a notification must be sent to the Business Owner to ensure that the Model is kept up to date, thus allowing for validation of the elements being recorded in the Data Lake or Database, Business Policy validation as well as the accurate representation of the data quickly. The Master Data Management solution should be able to seamlessly integrate into the business workflow while providing a “Management by Exception” view of the business for the appropriate Power User. This allows the Power User to validate, add, and correct any missing elements of the Master Data Model with the least amount of indirect or direct costs incurred.

7. Best Practices for MDM Implementation

Master Data Management (MDM) is a concept whose time has come. Enterprises that are just embarking on a MDM initiative would do well to look at examples already in place, in order to prepare themselves for the difficult road ahead. In this vignette, I will mention an MDM war story, and then describe three different best practices that can be harnessed for a successful MDM implementation.

MDM War Story A major Fortune 100 company makes the bulk of its revenue in a single quarter of the year. For that quarter, they have a MDM solution that dictates thousands of product and customer records, submitted by wholesalers and retailers. This MDM solution is so successful that it has run out of capacity for use by the enterprise. New wholesalers and retailers complain about the difficulty of getting business with an MDM that is too rigid; it will not allow wholesalers or retailers to access the product and customer information preprocessing. MDM is in the unenviable place of having to accommodate both the operational needs of the organization and the commercial needs of the business. And as importantly, the two sets of needs are not aligned. What MDM is seeking to do to build business value in the short term, is at the same time being eroded by attempts to leverage MDM to meet the operational needs of the enterprise during the non-peak part of the year.

Stakeholder Engagement As mentioned previously, MDM has to be something that every part of the business pushes forward and cultivates. All stakeholders need to collaborate, for example, marketing, finance, sales, logistics, and IT. Engaging with these stakeholders will ensure that everyone is invested in producing clean data for MDM. Some examples of engagement strategies include the following: Participate in both business and IT process sessions and workshops to identify all stakeholders and their processes, especially the interdependencies between business processes. Help stakeholders identify, measure, and agree to KPIs for all business processes, which are supported by MDM; the KPIs should also relate to important data attributes; KPIs should be both internal and external. Use the KPIs to prioritize data product development – after all, if customer service is your highest concern, make sure that the data products supporting customer service are constantly improved. The MDM KPIs should also emphasize blending business and technical perspectives; business stakeholders should be accountable for data quality from a business perspective, but IT is responsible for the costs of using poor quality data.

7.1. Stakeholder Engagement

An MDM project requires considerable investment by an organization, either in terms of technology and people, or in terms of risk exposure. Because MDM maturation involves changes to business processes, organizational roles and responsibilities, and in some cases a dramatic change in corporate culture, users must be evangelized and engaged throughout the implementation process. User buy-in can be sustained through iterative development, recognizing organizational milestones, and extending the sense of achievement through communication and recognition of contributions along the way. Stakeholder engagement is a critical component of any MDM program. Support from senior executives will provide the authority and resources necessary to define, implement, and govern an MDM solution. Business unit users with an interest in the quality and fitness-for-use of the master data need to drive the definition of data quality metrics and acceptable thresholds, information and business rules that determine the data content, and approve mappings to external or reference metadata sources. Information architects drive the technical mapping of data models used by the IT development staff to build interfaces or extract, transform, and load processes needed to feed the MDM solution. Business unit users are ultimately responsible for the quality of the master data. They set the quality metrics and thresholds, understand the data and business rules, identify masters vs. replicas, and are responsible for the use, enrichment, and ongoing validation of the master data. These stakeholders are also responsible for business process changes and training of other users who create and modify the data in business applications that consume MDM.

7.2. Incremental Implementation

Given the unprecedented momentum of MDM as evidenced by market acceptance and MDM vendor viability, there exists a risk for enterprises to embark on sweeping MDM initiatives that ultimately fail, either because change does not occur or the solution is too hard to deploy. To avoid these extreme possibilities, enterprises should consider beginning with modest, incremental MDM deployments that demonstrate stakeholder value; and if successful, build upon these initial deployments. An incremental approach reduces the risk of large-scale failure. When enterprise business users begin seeing business benefits from initial MDM deployments, it engenders higher levels of trust and ultimately increases the chances of getting the required support and funding for a full-blown enterprise MDM deployment. The key is to begin with an initial implementation that brings real business improvement for an identified business problem or process area within a collocated work group. MDM is about data quality – improving the quality of business decisions. The focus of the initial implementation should be on data used to support a significant business decision process.

Next, the initial implementation should target a practical, near-term-to-implement data quality business problem. Within this simplified objective, consideration should be given to one of the structural aspects of data: the data standard, the redundancies of record, and the relationship of data-specific entities. Each of these data attributes expresses a possible structural error in the data and there will be certain business processes or domains which are more amenable to an initial implementation focusing on just one of them. By design, the near-term data quality problem must be major enough from the perspective of a business process that operatively or strategically impacts the competitive position of the enterprise. Secondly, the identified initial implementation problem must not represent an organizational change that is too radical relative to the existing knowledge by business users of the problems and task.

7.3. Continuous Improvement

Continuous improvement involves frequent additions and enhancements to MDM to enable a more comprehensive set of solutions. The architecture design enables a continuous improvement roadmap to focus future projects or deliverables. Addressing more domains or business processes increases the business benefits enabled through MDM. Close collaboration with a governance board responsible for the MDM program supports frequent project changeovers and adjustments to missions and funding, thereby maximizing the business value.

The highest return on investment is from those projects that are business-driven and aligned with defined business strategies and also ones that are quick to deliver value based on capabilities. The governance process balances the technologies for prioritizing, funding, monitoring, and concluding projects within the roadmap. Efforts within the roadmap may contribute in relative degrees to operations, compliance, and innovation. Concrete areas of ongoing focus include: improving compliance processes using MDM; creating an MDM component architecture with base-case components, domain and application specifics; enhancement templates for domain and application components; and continual improvement cycles for pure MDM capabilities and components. These complementary initiatives ensure optimum use of current MDM capabilities while incrementally getting better and deeper enterprise help.

Some organizations go through more difficult and painful MDM method evolution on their way to operational stability and extensional success and longevity than others. But each approach, either fast or slow, is valuable. Each organization, although under-experience and under-preparation regarding one-stop, permanent, enterprise-wide MDM as envisioned, gains from living through the cycle to document and maintain an approach model to maximize advances toward that vision.

8. Conclusion

Master Data Management (MDM) addresses the problem of data quality and ensures the maintenance of trustworthy business meanings of shared enterprise data; it has, as one of its goals, the support for the interoperability of enterprise information systems. Digital enterprises emerge from the integration and interoperability of heterogeneous digital infrastructures, and MDM is an important enabler to ensure that enterprise information systems based on diverse and localized digital infrastructures provide trustworthy, relevant, and timely information for enterprise stakeholders. However, MDM has not been conceived as a complete and integrated solution but rather as a set of localized technologies. Moreover, MDM has not achieved the maturity levels of enterprise frameworks, data warehouses, and business intelligence; hence, it lacks the conceptual models, architectural views, best practices, and market acceptance developed by these solutions over a number of years.

In this essay, we have presented architectural frameworks for MDM that aim to provide those models, views, practices, and acceptance levels for MDM. We do not claim that our frameworks represent the ultimate models for MDM. Rather, we consider them significant progress on the current state of the art, which we hope will stimulate the development and deployment of more complete and integrated MDM solutions by industry practitioners and researchers. There are a number of factors that should stimulate the evolution to integrated solutions. The first factor is that MDM supports enterprise-wide processes that should, of course, be built upon enterprise-wide data. A second factor is that MDM solutions manage domain models that contain domains exclusively defined for MDM purposes. On the contrary, enterprise-wide and MDM domain models need to be built upon shared enterprise meta-models as in Data Warehousing and are then used by MDM solutions for MI services and for business applications to use MDM results. These shared enterprise meta-models are repositories that contain the semantic definitions of enterprise model elements. The semantic definitions contained in shared enterprise meta-models are extended during the MDM process in order to guarantee that the meanings associated with these shared meta-model elements are reliable over time, thus ensuring that the business applications of the enterprise use the trustworthy meanings of domain model elements.

8.1. Future Trends

Master Data Management (MDM) has come to be one of the work domains where most significant decisions are being made. So what is the future for this area of specialization? What trends could be decisive in the 'building of the future' in a sense? There are several thoughts to mention in this context. Together with Business Process Management and Information Architecture, MDM is likely to play a crucial part in the alignment of business and IT. MDM is the right source of the public value of the enterprise: it provides a stable for enterprise evolution, in a sense that all enterprise changes need to be correctly mapped / reflected / mirrored into MDM. MDM solutions should progress towards allowing for change requests to be efficiently and collaboratively evaluated in order to estimate their real impact on the enterprise.

Both business and technical stakeholders interpret enterprise changes in a different way, with a different granularity level, and without sufficient formal constraints bridging the gap between their respective representations. Requests originating from business can be translated into corresponding change requests for related enterprise cultural, organizational or functional aspects, both formalized, classified, and associated to a categorization model of the enterprise. So, within the near future, change request management should become a new crucial component of MDM solutions. Currently not envisioned, this requirement appears to be mandatory for the global enterprise because of its prior beneficial impact on the improvements of transparency, availability, quality, and efficiency of the core enterprise processes and of the services provided.

9. References

1. Motamary, S. (2021). Machine Learning Applications in Demand Forecasting and Order Fulfillment for Smart Manufacturing OSS. International Journal of Science and Research (IJSR), 10(12), 1639-1653. https://www.ijsr.net/getabstract.php?paperid=MS2112143019 https://www.doi.org/10.21275/MS2112143019
2. Koppolu, H. K. R. (2021). Data-Driven Strategies for Optimizing Customer Journeys Across Telecom and Healthcare Industries. International Journal Of Engineering And Computer Science, 10(12).
3. Yellanki, S. K. (2021). Operational Efficiency through Service Integration: Insights from Platform-Based Business Models. International Journal of Engineering and Computer Science, 10(12), 25756–25772. https://doi.org/10.18535/ijecs.v10i12.4681
4. [4] Pandiri, L. (2021). Cloud-Based AI Systems for Real-Time Underwriting in Recreational and Property Insurance. International Journal of Science and Research (IJSR), 10(12), 1626-1638. https://www.ijsr.net/getabstract.php?paperid=MS2112142733 https://www.doi.org/10.21275/MS2112142733
5. Kalisetty, S. (2018). Cloud-Driven Big Data Harmonization for Real-Time Demand Forecasting in Hybrid Retail-Manufacturing Supply Chains. Global Research Development (GRD) ISSN: 2455-5703, 3(12).
6. Gadi, A. L. (2021). The Future of Automotive Mobility: Integrating Cloud-Based Connected Services for Sustainable and Autonomous Transportation. International Journal on Recent and Innovation Trends in Computing and Communication, 9(12), 179-187.
7. AI-Based Financial Advisory Systems: Revolutionizing Personalized Investment Strategies. (2021). International Journal of Engineering and Computer Science, 10(12). https://doi.org/10.18535/ijecs.v10i12.4655
8. Dwaraka Nath Kummari. (2021). Smart Infrastructure Auditing: Integrating AI to Streamline Manufacturing Compliance Processes. Journal of International Crisis and Risk Communication Research , 168–193. https://doi.org/10.63278/jicrcr.vi.3036
9. Machine Learning Integration in Semiconductor Research and Manufacturing Pipelines. (2021). IJARCCE, 10(12). https://doi.org/10.17148/ijarcce.2021.101274
10. Digital Infrastructure for Predictive Inventory Management in Retail Using Machine Learning. (2021). IJARCCE, 10(12). https://doi.org/10.17148/ijarcce.2021.101276
11. Meda, R. (2021). Machine Learning-Based Color Recommendation Engines for Enhanced Customer Personalization. Journal of International Crisis and Risk Communication Research, 124-140.
12. Dwaraka Nath Kummari. (2021). A Framework for Risk-Based Auditing in Intelligent Manufacturing Infrastructures. International Journal on Recent and Innovation Trends in Computing and Communication, 9(12), 245–262. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/11616
13. Kalisetty, S. (2017). Dynamic Edge-Cloud Orchestration for Predictive Supply Chain Resilience in Smart Retail Manufacturing Networks. Global Research Development (GRD) ISSN: 2455-5703, 2(12).