Panamerican Mathematical Journal
ISSN: 1064-9735
Vol 34 No. 4 (2024)

Intelligent MDM Systems: Using Agentic AI and Machine Learning to Harmonize Customer and Product Master Data Across Financial Domains

Ramesh Inala,

Data Engineer,ORCID ID: 0009-0009-2933-4411

Article History:
Received: 18-09-2024
Revised: 24-11-2024
Accepted:04-12-2024

***Abstract**:* Master Data Management (MDM) aims to ensure a harmonization of business-relevant master data entities showing consistency, accuracy, and synchronization throughout different systems and processes. While companies are able to a certain extent to implement their own MDM systems leveraging capabilities of the prevailing Enterprise Resource Planning solutions, these systems tend to be not enough to master the challenges of controlling and utilizing Master Data. Additionally to the MDM capability of traditional ERP solutions, intelligent MDM systems are required that are purpose-built yet increasingly integrated within the overall technology stack. Such intelligent MDM systems harmonize customer master data and product master data, enabling a 360-degree view of business partners' transactions,, and products sold, while also providing a single source of truth to support transactional and analytical capabilities of different enterprise applications and data lakes. Harmonizing customer and product master data is the overall focus of our paper, discussing the function, motivation, challenges, and limitations of intelligent MDM systems.

As companies disentangle their technology architecture due to the need for faster innovation cycles and more agile organizational structures, and because a wider network of partnership ecosystems becomes important, the demand for MDM capabilities will further increase. However, as master data are the essence of enterprise applications as well as the usage of business- and transnational data analytics and artificial intelligence, the need for intelligent MDM capabilities goes beyond merely improving the capabilities already offered by existing ERP solutions. Therefore, we elaborate on what we determine as intelligent MDM systems and discuss how these systems contribute to customer-centric and product-centric companies. To supplement insights from an extensive literature review on MDM systems, we conduct semi-structured interviews with around 20 experts from the areas of MDM, enterprise applications, enterprise data, and business analytics.

Keywords: Intelligent MDM, Master Data Management, Enterprise Applications, ERP Integration, Data Harmonization, Customer Master Data, Product Master Data, 360-Degree View, Single Source of Truth, Data Synchronization, Data Accuracy, Business Analytics, Technology Architecture, Agile Organizations, Partnership Ecosystems, Data Lakes, AI Integration, Transactional Data, Analytical Capabilities, Customer-Centric Strateg

Introduction

In order to enable rapid and accurate decision-making, all companies need to rely on their Intelligent Master Data Management (MDM) Systems to provide information about their business environment, and then transform that information into the knowledge that underlies such decisions. An MDM System is intelligent to the degree that its functionalities automate intellectual tasks, i.e. are not an end in itself but a means to speed up decision processes and decision making. In the face of Big Data which demands near real-time decision-making processes, intelligence is mainly delivered through the provision of context-relevant suggestions, notifications, and alerts, and by steering the user experience towards intelligent paths of exploration and decision-making. MDM is a general term that covers all information organization processes that allow a company to set up, manage, and maintain a master data structure indicating especially the characteristics of the most important objects related to the company's business objectives. As such, it includes activities such as establishing a Master Data Dictionary, providing the data for and maintaining an organization chart, as well as establishing product and customer master data. Among these, the most important and representative of MDM as a discipline is customer and product master data management, as those two entities are the primary objects to which the company necessarily pays its attention. The establishment of these two entities is the necessary pre-condition to successfully carry out any other activity in the company's value network.

1.1. Overview of Key Concepts in Master Data Management

Master Data Management (MDM) provides a way to manage the critical information of an organization — the master data — consistently and accurately throughout a data ecosystem that is becoming increasingly complex and connected. A coherent and harmonized view of master data provides organizations with a golden record of truth that is necessary for analytics, decision-making, customer- and partner-oriented operations, and compliance. Customers face organizations that appear chaotic, uncoordinated, and inefficient when basic data exchange and options are not managed appropriately. Organizations face customers who seem jumpy and are increasingly willing to switch their favorite supplier or service provider when a more attractive option is available.

In particular, Product Master Data (PRMD) and Customer Master Data (CUSTMD) are key components of the organizational master data set. Proper management of these components in terms of quality, consistency, and timeliness is required if organizations want to use technological advances in the area of data analytics and data sharing and exchange. These advances contribute to the increasingly connected and coordinated character of logistics and value chains in the operations of organizations, and hence also to the customer- and partner-oriented component of the corporate mission. Harmonization of the two data sets provides organizations with the necessary basic information to understand relationships with customers or suppliers and to analyze corporate performance from a data-driven perspective. MDM provides the ability to move from anecdotal evidence to fact-based and quantitative decision-making in all organizational functions. Using opportunities related to the information economy requires organizations to work in a data-driven decision-making mode, where facts — and reliable master data are essential ingredients for those facts — are guiding the strategic and operational decisions made.

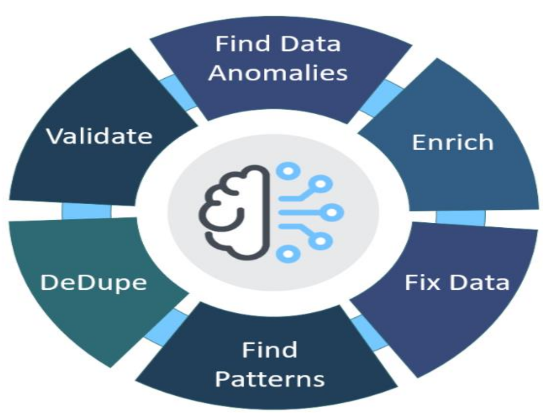

Fig 1 : Agentic Customer Master Data Management (CMDM)

2. Understanding Master Data Management (MDM)

Data is an enterprise's most precious resource. It is the fuel for making good business decisions. It guides managers in improving business performance. Deficient data leads to questionable decisions, wasted effort, and potentially adverse consequences. Master data is a subset of all organizational data that is used several times by multiple applications or systems, which is why it needs to be managed effectively. For this essay, a simple definition of what Master Data might be: Master Data is a consistent and uniform set of identifiers and extended attributes that describe the core entities of the enterprise, which usually include the Customer, the Product, the Employee, the Vendor, the Brand, the Location, and the Asset. Master Data Management (MDM) is the process of enabling an organization to singularly, accurately, and persistently define and manage its critical data entities.

MDM solutions are designed to support this capability so that there is a single authoritative reference source for all business-critical entities. MDM Systems are the enablers of MDM. These are tools that create a single Master Data Repository for any or all of the enterprise's data domains, populate the unique trusted record per domain, and distribute the information to all internal and external applications and participants who need it. MDM is an organizational process supported by policies and procedures, which are generally defined by the Business Users of Master Data Managed Systems or Data Stewards. The goals of MDM are usually defined by the organization's Architects. MDM has a team of Data Architects, Solutions Architects, Project Managers, and Subject Matter Experts to achieve these goals. They work with Business and IT sponsors to create road maps and develop business cases, as well as approve the value achieved by the successful implementation of MDM Projects.

quation 1 : Cross-Domain Data Harmonization

$$\mathcal{H} = \frac{1}{N} \sum_{i=1}^{N} \left(1 - \frac{||x_i^{(d_1)} - x_i^{(d_2)}||}{||x_i^{(d_1)}|| + ||x_i^{(d_2)}||} \right)$$

$\mathcal{H}$ = Harmonization score

$x_i^{(d_1)}$, $x_i^{(d_2)}$ = Features of entity i in domains

N = Number of matched entities

2.1. Key Principles of Master Data Management

MDM relies on a number of principles that distinguish it from other related fields. So, what is, specifically, MDM? MDM is concerned with the life cycle, the management, and the governance of master data. The life cycle comprises the aspects of creating, updating, and deleting master data. It connects MDM to disciplines such as data modeling and the construction of data warehouse dimension tables, and it couples it with DQ activities such as data profiling and data quality stewardship. Ensuring data quality is certainly an important aspect of MDM and there is a close cooperation and interlocking relationship with DQ.

Master data management is also concerned with data governance - establishing accountability, policies, and rules regarding master data. Yet, it is not about establishing rules for all company data, and not all rules are MDM-specific. Data governance can best be understood as collaborative oversight where business and IT together establish policies and rules regarding master data. Thus, MDM can be regarded as the core of a data governance program. MDM can also be seen as a state that is achieved through a set of MDM processes and the MDM store, usually a kind of data hub.

MDM creates a reference point for operational systems using master data. These operational systems refer to the golden records of the MDM store where they can obtain a consistent view of the master data. In a harmonized way, initialized by the appropriate MDM processes, the golden records are then made available to the operational systems for day-to-day business to follow their business processes.

3. The Role of Agentic AI in MDM

Data-driven smart decision-making is the fuel governance systems, private and social, including governments, banks, businesses, distributive services, or any other type of authority or organization, increasingly need to inform their actions. The volume and variety of transactional data accumulated by these systems are enormous and surpass any person or group manpower to validate, control, and conceptualize it before its final form, and before unleashing its theoretic potential, in useful knowledge that allows for monitoring and improving the normal state of affairs and prepare for impactful change. Therefore, for knowledge to come at the right time and benefit the largest number of people, the use of intelligent systems is indispensable. The consequent demands for the development of an Intelligent Master Data Management System in charge not only of master data validation but also of systematization and actualization of the stored knowledge are both natural and dramatic. This paper emphasizes the role that intelligent decision-making systems, or Momentum Agentic Intelligent Systems, whose concept is presented in the first section, whose core is Agentic Intelligent Systems and named after its main driver, i.e., the Momentums, or moment-to-moment intelligent decision-making, should have in the construction of Intelligent Master Data Management Systems. For that, we first present MDM and the process of how MDM systems prepare master data to be stored by knowledge accumulator data repositories and the Agentic Intelligent System basic concepts implied in the construction of intelligent systems and possible use of Agentic Intelligent Systems as a support to the building of Intelligent MDM Systems for knowledge control. Then, we summarize the most relevant characteristics of MDM systems. In this respect, we show two MDM projects stressing their MDM features and discuss MDM importance on how gravity has sometimes been neglected during the building of data platforms. This summary then brings to light some of the gaps existing when MDM solutions are compared to traditional industrial rigors, most of them guided by data accuracy, and reviews how Agentic Intelligent Systems could reduce the gap without imposing high hurdles to MDM solutions.

3.1. The Influence of Agentic AI on MDM Implementation

Master data management (MDM) was invented with the aim of storing the most relevant business data of any enterprise on an internal cross-corporate system that could be utilized for all enterprise applications while still allowing the important external connections of customer and supplier

Panamerican Mathematical Journal
ISSN: 1064-9735
Vol 34 No. 4 (2024)

information available through additional channels. A current limitation of those systems is that they need to be updated manually or with automated data feeds, while capturing only limited information, such as the country and the line of business for enterprise customers. However today, more often than not, enterprise MDM systems lack the very timely enriched external information available from interconnected AI models.

In the AI era, these limitations may be fundamentally changed, and not just because of systems that enrich the internal MDM systems with AI chat capabilities. Instead, the capabilities of agentic systems leveraging knowledge models and decision-making or data validation models could become important drivers for the business model as well as the implementation of MDM. For example, any service company that enhances the internal MDM with the ability to engage in more interactive dialogues with their customers and suppliers for internal data quality reasons or to improve customer experience could well differentiate themselves slowly vis-a-vis competitors that are not taking that step. The same strategy could be pursued by a B2B eCommerce company feeding their MDM with the transaction data, since shipping, info requests, and other logistics questions could all be modeled on such a dialogue-based interaction.

Fig 2 : (MCP) Server to its Agentic MDM to Power Autonomous

4. Machine Learning Techniques in MDM

Master data management (MDM) is a complex task that regularly requires the action of a workforce to incrementally correct data. For this reason, the work is of high complexity and can have a higher cost than expected due to the inherent errors in MDM tasks, especially for huge datasets. For example, entity linking in product master data with thousands or millions of products is hard and costly. Given the need for solutions that can be put into action to support this human-in-the-loop work needed to manage data quality and consistently improve the performance of MDM processes, many machine learning techniques have been proposed to address some areas of MDM like deduplication, entity resolution, fingerprinting, data enrichment, data auditing, and so on. Classically, the proposed approaches in MDM have been traditionally focused on the integration of data from multiple sources to create a common view to manage. Some of the main derived challenges from data integration are similar to MDM challenges, thus ML approaches already applied in data integration are also applicable to MDM tasks and vice versa.

In particular, these approaches are focused on two particular areas of MDM, data quality and data enrichment. The data quality domain contains some tasks embedded in continuous data management, like corrections, monitoring, and auditing. The data enrichment domain refers to the process of obtaining additional information coming from different data sources about the master data. In summary, given that our vision is to build an intelligent and harmonized MDM, exploring the use of supervised, unsupervised, and reinforcement learning approaches seems to be a good way to go, since they can contribute not only to the improvement of data quality itself but also boost tasks like data curation addressed to enhance the degree of consistency of the data available to the organization.

4.1. Supervised Learning Approaches

In supervised training, the task is usually to predict labels from provided feature values for the labeled input data. It is indeed a classification task when labels are discrete. Supervised algorithms can also be used in regression tasks to predict numerical target values. In MDM, the labels can usually be identified by the expert either due to labeling or due to pre-defined rules. The provided properties can be either a subset of the existing properties, which were already to some previous extent analyzed or any conclusive set of properties. In tuple functions, tuple-based algorithms like classification trees or logistic regression are favorable to classify single data tuples as function attribute values. In pair functions, relationship-based algorithms like pairwise classification with suitable classifiers are favorable to predicting the values of the pair attribute. Pair functions can be either applied alone regarding only one of the two attributes or be derived reciprocally based on one pair function regarding each attribute. With respect to the considered MDM tasks, we classify the content of the supervised training techniques into four categories. Thus, it summarizes supervised learning techniques for MDM in a compact visual. It provides an overview of how supervised techniques are applied to MDM tasks. It provides detailed concepts of how the MDM tasks are performed and how the embedding is done for the supervised techniques. In supervised tasks, a predictor trained on previous observations of records is built to derive missing values for tuples that have missing values in certain attributes. These problems are common regarding several aspects of MDM tasks like the entity resolution task of detecting duplicate records.

4.2. Unsupervised Learning Approaches

An important class of machine learning techniques that can be applied to MDM solve the problems of the absence of labeled data describing the different classes in such data. In particular, MDM needs to use a huge amount of unlabeled data, which can be used to learn how to relate and harmonize the customer and product master data of a company. For this reason, unsupervised learning can be one of the main techniques used by MDM. In practice, many of the solutions that build an MDM component use some unsupervised learning technique almost as a required first step for defining the operations that are supported by each MDM component.

For example, the entity resolution component must rapidly find the pairs of records that probably contain the same real-world entity among the millions distributed in different customer and product reference data stores in the organization. This task is performed by detecting duplicate records in such stores without any prior information on the keys that make the two record pairs be considered as representing the same real-world entity; this process is usually called address validation. Also, the merging component must transform an unordered collection of records of the same entity into a single comprehensive record. While graphical modeling techniques can be used to restrict the number of feasible solutions to the two previous tasks, the only technique to generate candidate record pairs used in most current practical systems is data clustering. Therefore, a clustering method must define what data similarities, and in particular what distance metric, have to be used and what the predicted number of clusters in the two unsupervised learning problems formulated by these two MDM components. Unfortunately, supervised learning cannot help answer these questions. These components therefore cannot work without any information previously extracted from the data.

4.3. Reinforcement Learning Applications

The approaches so far presented are sufficient for every kind of data analysis task. They are focused on learning data patterns or associations or modeling their data distributions. However, MDM also faces the problem of decision-making. Given a specific type of data, specific actions must be taken, for example; categorize, clean, or deduplicate. These decisions can often take a huge amount of time, which will result in higher labor costs for a company.

Reinforcement learning is a strong candidate for solving the domain of decision-making and can receive large datasets as input. It addresses the problem of how an agent can automate the decision implementation process based on feedback signals received from the environment. MDM in the data preparation and data editing phase has to face the process of lacking or incomplete data in the attribute values or possible mislabeled categorical attributes. The feedback signal is in most cases a specialized observational function implemented in MDM. It can measure the classification accuracy, the accuracy of attribute correlation analysis, or data cleaning results. The reward signal is then based on which decision will yield the most improvement and how much it improves. At this point, there are no methods explicitly designed for MDM that adapt reinforcement learning but only fundamentally generic applications that utilize reinforcement learning.

Another example is the combination of title prediction, title cleaning, and title deduplication in online bibliographic data. User studies show that the costs of cleanup can be minimized by predicting which titles are likely wrong and should be inspected by editors. Therefore, an estimator is used to assess which types of data errors exist, and the editor is reformulated as a reinforcement learning agent that automatically decides the optimal action stored in a value function, avoiding editing the title for edit summaries with distributions that are statistically similar to those produced by humans on comparable tasks.

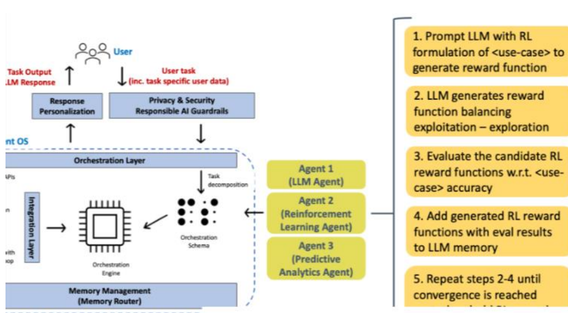

Fig 3 : Reinforcement Learning Agents for Industrial Control Systems

5. Financial Domains Overview

We illustrate the importance of Financial Domains for business in general and MDM work in particular, by giving a detailed overview of 4 different subdomains in Finance: Banking Sector, Insurance Sector, Investment Management, and Fintech Innovations.

Whereas Knowledge is Money as a synonym for our knowledge economy, we can summarize the Financial Domain by stating that Money is Money. The whole purpose of the Finance is to make money. All other sectors and domains generate added customer value with their products and services, which, in turn, creates a demand for money. The Financial Domain, MDM Services included, makes no additional contribution to society, only serves to take their cut, to 'profit' without a loss, but indirectly deciding which products and services are further demanded, wanted or not, used or not. MDM Services assist businesses in increasing their margin but have no impact on revenue.

The Banking Sector includes all MDM Applications related to financial services provided by commercial banks or investment banks. Banking includes business on deposit and close-to-deposit accounts, loans, credit and debit cards, as well as payment transfers, how they operate, and which kind of credit their customers ask for. The Insurance Sector constitutes all MDM Applications dealing with property and life insurance policies, their customers, and the risks they cover. Investment Management covers MDM Applications related to portfolio management, organized in assets, listed or not, and their multifaceted ownership structures. Nevertheless, all investment vehicles placed by customers in custody by Banks are also partly covered. Fintech Innovations represent the disruptive technological novelties and developments primarily initiated and operated by Fintech companies, using Distributed Ledger Technology, Peer-to-Peer Platforms, Alternative Credit Scoring, Mobile Banking, Machine Learning, Automation, Data Investment, as well as Artificial Intelligence.

Equation 2 : Agentic Data Resolution Confidence

$$\mathcal{C}_{\text{res}} = \sigma\left(\theta^\top \cdot [\text{ML}_{\text{match}}, \text{AI}_{\text{intent}}, \text{Rel}_{\text{net}}]\right)$$

$\mathcal{C}{\text{res}}$ = *Resolution confidence*

ML{\text{match}}$ = Similarity score

$\text{AI}*{\text{intent}}$ = Agentic intent factor*

*$\text{Rel}*{\text{net}}$ = Network trust score

θ = Learnable weights

σ = Sigmoid function

5.1. Banking Sector

The foundations of the banking sector lie in Financial Statement Accounting, providing information on how organizations possess and utilize assets while also ensuring an understanding of relations with third parties. Banks serve as a platform for exchanging financial products, utilizing deposits to engage in lending, invest in securities, and provide related services at a cost.

Banks primarily function as liquid funds and generate resources that are mobilized for capital formation in the economy. They act as custodians of the savings of the people in a country and also work on generating economic resources that can be effectively utilized to achieve economic development and to establish a correlated distribution of resources. In this respect, Banks perform a unique but risky function in the economy by issuing transaction services that enable individuals to save on the costs of transacting. Despite the often-unique risk profiling of different client types as well as product themes, Banks develop crucial product-service combinations from fees based on earned income as cost structures that are predominantly liquidity and interest rate driven.

Proper visibility is provided to Investors, customers, and counter services through Bank Financial Statement Communication. Financial Corporate Banking, Bank Private Banking, and Investment Options act as the services provided for Investors. Retail Banking Services, Depository Services, and Specialized Services act as the services provided for clients. Bank Accounting acts as the Basic Infrastructure supporting the Financial Services. Bank Financial Services Data Production, Bank Planning, and Control and Bank Financial Service Products act as Framework Services. Like Manufacturing, trading, and Financial Portfolio Ownership Corporate Industries, Banks also follow Financial Statement Accounting. Bank Asset and Liability and the Income Expenditure Statement act as Core Financial Statements.

5.2. Insurance Sector

The insurance sector offers a considerable variety of products for covering risks provided by a wide range of companies through many sales channels. In insurance, loss events are usually unforeseen and therefore cannot be planned. For managing the financial engagements that arise from a claimed insurance event, insurance companies use customer and product data. Policy data and coverage data are key to describing these data.

In the area of policy data, the distinction has to be made between core data (such as information on coverage) and detailed local data (e.g., information on the quality of the object that is insured). The latter is organized by postal code and is essential for calculating the expected value of the events and thereby the reliable planning of premium payments over the years until there is an event. For agreed contracts, long periods of time between payment of the premium and the provided service exist. In addition to core policy data, adequate financial planning for settlement periods that may not be foreseeable is needed. In the area of coverage data, the focus is on products; that is, contracts for covering risks which are defined by detailed information about in-scope and out-of-scope risks. Additionally, some life insurance companies offer products that can be transferred to partners or invested for the lifetime of a customer.

5.3. Investment Management

Investment management is both an art and a science, combining an intuitive understanding of markets and traders’ psychology with sophisticated statistical mathematics and modeling techniques. The service is often offered by specialized firms and wealth management divisions of banks to moneyed individuals, families, and institutions that need support in growing their wealth, or managing their savings while living off the investment income, in the presence of various constraints. Investment managers create profit for clients by making market timing bets, security selection decisions, or both. Real investment portfolios rarely consist exclusively of liquid exchange-traded funds and assets, and thus the trade-off between expected return and risk is not a straight line. Due to the constraints, including the need to hedge risk arising from the client's declared intentions, investment management products are quite diverse and also tend to be implemented as strategies rather than as distinct products.

Investment management requires an understanding of many disciplines and is distinct from the function of merely executing trades, or determining optimal market-clearing trade prices based on high-level econometric models. This discipline spans making predictions on the behavior of physical entities, utilizing tools from various disciplines—engineering, programming and information systems, economics and political science, and research—to evaluate the corresponding real assets and flows, and then actively or passively timetabling the necessary trades on the stock, bond, commodities, or real estate exchanges.

5.4. Fintech Innovations

Financial technology or fintech consists of a range of tech-driven innovations applicable to financial services. The word fintech first appeared in the 1980s when it designated only the back-end technologies available to banks and stock brokerages. The term has since evolved to apply to the assortment of innovations, products, and solutions emerging in the financial services sector and offered to businesses and consumers alike. The modern sense of fintech first emerged in the wake of consumer services, such as online money transfers, which became possible thanks to the commercialization of the Internet around 1995. The sector grew quickly over the next ten years and exploded after the financial crisis of 2007–2008 when financial hedge funds and banks began offering services to businesses as well as using these new technologies to bolster efficiency on the ground, drive down costs, cut paper trail, faster execution, new markets, and new trades, and threaten established players with new systems and solutions that could do it quicker and better.

The fintech world exploded once more after the internet became mobile, ushering in new mobile-only banks, solutions on both the consumer and the business side, and omnichannel or integrated business banks that could handle both aspects at once. Some of the products available for consumers include social transfers, payments and exchanges, borrowing, saving and accounts, investing, insurance, and even accounting. The services available to businesses run the gamut of back-end accounting and enterprise resource planning to customer relationship management, expenses, financing, invoicing and payments, investment, payroll and HR, risk management, and payments and collections. The increasing sophistication and pace of fintech innovation means that omnichannel banks, those focused on individuals, and diversified bank-like firms will have to integrate their full range of offerings for customers.

Fig 4 : Benefits of Using AI in FinTech

6. Challenges in Data Harmonization

Harmonizing product and customer domain data creates new challenges when compared to regular domain data consolidation tasks. Perhaps the most important harmonization challenges arise from traditional operational master data management practices. These practices are managed in decentralized user departments while data in the domains should be user-department-independent. This boils down to two significant challenges: different quality notions and domain-specific taxonomy definitions.

1. Data Quality Issues

Data quality issues are well-known in data integration tasks. Traditional master data stores have been developed by the different user departments often for their purposes. Quality of product words is therefore not a shared responsibility, governed by a global quality profile. Different quality notions exist, based on the fact that the quality of the data has only been assessed from the perspective of the local department; different department-specific business rules apply. In marketing scenarios, the quality of customer profiles is even taken into account for computing ranking algorithms. This influences the tedious travel prediction algorithms, but can also be useful for data mining purposes. Quality scores assigned to the data are stored in the profile dimensions of the customer dimension.

2. Integration Complexity

Integration complexity comes in more than 1 variant when consumer and product data are merged. The first concerns the relations captured in the taxonomy. Product dimensions are usually mapped to taxonomical classification systems of different product vendors. The resulting taxonomy organization is vendor-specific; the vendor's taxonomy organization is contained in the relationship schema of the product dimension. No integration of the different taxonomy organizations is required by the quality notion, assuming that every product domain is served by a different vendor. On the contrary, the customer space taxonomy is user department independent, and must therefore be organized according to general semantic linguistic criteria.

6.1. Data Quality Issues

Master data in an MDM system serves as a standard by which transactional data can be validated, cleansed, transformed, and standardized. Organizations rely on an MDM system to help achieve a gold-standard view of their customers and products, and this fact implies that the data in their MDM system must be of very high quality. Data quality, however, is a challenge in many organizations, and enterprises become wary of implementing an MDM system due to worries that the information in it will not be high quality enough to support transactional processing in other enterprise systems. Therefore, the first challenge to overcome when implementing an MDM system, especially for PIM is to ensure that the quality of the input data is sufficiently high, else there is little value in enabling a gold standard of product data to be created. Implementing MDM for PIM in particular also holds the risk of creating a central repository of product data that becomes the source of truth for reduced time-to-market and error rates for transactional data and that cripples the organization if the accuracy and timeliness of the MDM data falls. The adage of junk-in junk-out must be addressed and a high-value data governance strategy must be built around the MDM PIM solution.

The low priority given to PIM and third-party product data by organizations is usually because PIM activities generally incur an operational expenditure rather than create revenue-generating process improvements. However, with the recent explosion in the number of digital commerce channels, creating attributes that properly classify products increases the fixed cost in order to create the data once for each item category and then put into place efficient processes to maintain the quality of the data. The need to provide product data in various formats and categories appropriate for each channel and the operational efficiencies that come from having the MDM system provide relative comparisons of product data attributes at specific categories indicates the return on investment on MDM for PIM.

6.2. Integration Complexity

In an MDM system, the objectives of resolving entity identity discrepancies and ensuring that there is only one record for each customer, product, or service are relatively easy to achieve. However, the actual records that microservices use to run a business are often distributed across multiple source systems. To make MDM deployment a seamless part of enterprise workflow, MDM microservices must provide an integrated view of the entire enterprise environment; and this, in turn, requires the orchestration of multiple microservices which integrate data from various data domains. In the case of a single domain, this can sometimes be achieved using processes, but for the domains that concern customers, integration of data cannot be done quickly using classical processes. Furthermore, the involvement of multiple entities during the life cycle of a transaction increases the integration complexity.

Limited by the need for integration and the intrinsic shared nature of customers across functional domains, the C360 view construction cannot be limited to custodial or archival systems such as sales or customer service history and relies on the relatively infrequent need to access older transactions. Instead, dedicated microservices that are geared toward facilitating the integration of disparate transactions that involve a single customer, although they may span different functional domains and have been recorded during the processing of different business processes, need to be deployed. The processes that these microservices enable need to be very low-latency workflows. Although some form of integration has always taken place for the domain of products and services, this integration has normally been at the data level rather than transactional; it has relied on the assumption that the integration processes that were running at the data level were obtaining valid associations; and it has focused on association prediction rather than association maintenance.

6.3. Regulatory Compliance

New regulatory compliance mandates can emerge and can be identified in Trend Analysis, helping companies to take proactive action and be one step ahead of their competition. Rules and regulations require that customer and product data adhere to certain mandates. In the case of customers, anti-money laundering or KYC regulations mandate that sensitive data about customers be recorded and remain private. In the case of financial institutions, the entity in the US responsible for the enforcement of these money laundering laws is known as Customer Due Diligence (CDD). Money laundering mandates require that financial organizations collect sensitive particulars about their customers, including name/address verification, date of birth, and tax identification numbers. Violating KYC data guidelines can trigger fines along with other regulatory actions and consequences. Protecting customer data is of paramount importance in ensuring compliance. Therefore, companies will invest significantly in customer data compliance capabilities.

Similar and additional regulations apply to product data. For consumer products, compliance with numerous regulations covering labeling and recall is mandated. In the case of restricted consumable products, compliance with mandated and specified consumer food product labeling is required. In the case of pharmaceuticals, compliance for prescription and over-the-counter drugs that contain active ingredients, which can impact their functions, is mandated. Manufacturers are required to label their packaging to promote the safe use of those drugs. Failure to comply with Company Product Data mandates can expose manufacturers to significant fines and penalties. Product and customer data compliance checks are critical components of MDM tools. Capable systems must automatically verify regulatory compliance and monitor for changes across the lifecycle. Compliance data can be maintained to ensure current customer and product compliance.

7. Future Trends in MDM

As organizations embark on their digital transformation journeys fueled by emerging technologies, they encounter a paradox of increased agility, fluctuating customer expectations, and the demand for real-time decision-making based on verified trustworthy data. Business leaders must prepare their organizations and data infrastructure for this complexity. Businesses are exploring new ways to harmonize customer and product data while providing an exceptional experience to their customers and partners. The primary challenge faced by organizations is the management of various master data types, while also achieving data governance maturity with their MDM program. Digital transformation initiatives based on innovative technologies, such as artificial intelligence and machine learning, along with real-time internal and external data feeds, can facilitate advanced automation of the data governance processes.

MDM system users will expect intuitive experiences akin to well-designed, consumer-grade systems. These user interfaces will employ data visualization techniques, embedded intelligence, and chat interfaces to allow users to ask questions, locate the correct data, and understand the results. Because data management processes are more often governed via collaboration from many data stewards, and not only a few that are assigned to master data roles, an advanced social enterprise technology will allow more data users to foster trust and confidence in master data and to expect more predictable quality on the master data that they consume. Information stewards will be alerted when data quality anomalies occur, either through background monitoring of data quality assessments tied to specific data domains or when specific users need to accomplish a task in a non-compliant manner.

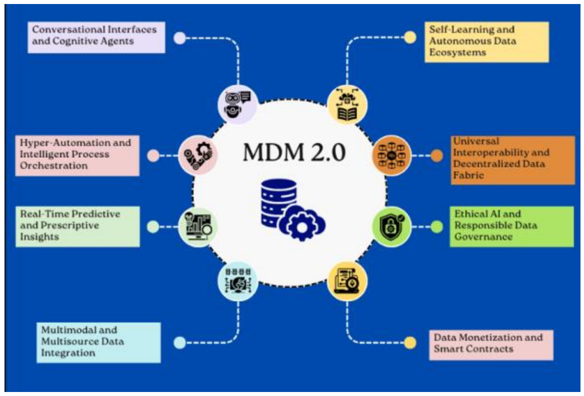

Fig 5 : Trends in Enterprise Data Management

7.1. AI-Driven Data Governance

The symbiotic relationship between artificial intelligence (AI) and data governance is poised to reshape the future landscape of data management, triggering a paradigm shift toward next-generation data governance practices. In the context of information technology (IT) and data management, governance refers to making sure that every piece of information following its lifecycle from cradle to grave is of high quality, is taken care of while being used, is available where it is supposed to be used, is mildly protected from unauthorized access, and is compliant with any regulatory requirements at each phase of its lifecycle. In a modern corporate environment, sensitive and critical data is constantly being exchanged and transformed in many different manners. This makes data governance a challenging endeavor for any organization, and AI might just be the right technology to help face this challenge.

AI technologies can make data governance easier, faster, and less expensive to execute. The use of AI enables the realization of a more automated and continuous model of data governance that can significantly reduce technical debt. This can be achieved through aspects such as intelligent automation of manual processes, smart suggestions for workflows and notifications, automated detection of issues and risks, alerts for risks that need immediate attention, intelligent recommendations on business rules and policies, intelligent tracking of data usage and violated policies/rules, intelligent detection of sensitive data and its usage patterns, and AI-powered end-user training, all of which can ultimately lead to more efficient data governance activities at the organization. By enabling organizations to process data faster and more effectively at greater volumes and to address more of their governance needs, AI moves data governance closer to the point of action, lowers the overall costs of exercising governance, provides the potential for greater intelligence from governance efforts, and allows data governance to support more corporate objectives than previously possible.

7.2. Real-Time Data Processing

The need for real-time data processing is driven by the strategic perspective on data and the industry mergers and acquisitions taking place. Companies realize that their wealth isn't stored in their products or balance sheets, but in their databases, and that the master data within ponds the essential customer and product edges. The business value related to those keys for external marketplaces is being measured in multiple million Euros. Companies specialize in acting as external data pools for someone else's business. They do business with product manufacturers and retailers who sometimes stumble in the negotiation process because they don't have the product edge data of the current deal level. Once established, the relationship triggers the publishing of promotional products matched against the product master data. The MDM product hub thus becomes an active participant in the data-driven enterprise economy.

In order to act as such a valuable resource, the data in the MDM hub must be current, reliable, and complete. This favors real-time data processing activities. Although many processes in classical MDM implementations were batch-driven, the sales and marketing functions of a company already demand quicker publication cycles for product data in connected marketplaces. The flow of product and product-related data will be energy-driven by customer data templates that allow automatic updating of the product information that is already present in the marketplace from a manufacturer’s templates. Validation is inevitable, but proactive approaches help here, too.

Data publication templates for partners in the supply chain will allow a continuous checking of the external data for synchronization with the internal product data and assist in the establishment of data exchange projects. Marketplaces of MDM systems will continue to deepen and offer basic data validity checks on imports. Software solutions processing event streams address the incoming flood of update requests. With software solutions, the data can be replicated for cloud or hybrid environments, which take over the front-end delivery functions. In such cloud scenarios, speed and flexibility become the core benefits.

7.3. Enhanced User Experiences

Traditional Master Data Management (MDM) has mostly focused on addressing the demands of data governance and data processing efficiency. Typical topics for MDM implementation have been compliance, data quality, treasuring "golden records", creating and maintaining the central data repository, providing reports, and utilizing data in Business Intelligence tools. With the advent of Artificial Intelligence and Machine Learning, business processes are becoming more flexible and user-driven. Additionally, new technology trends such as the attention economy and new user interfaces such as chats open new opportunities for data and analytics services supporting end users. This is driving new user experiences for MDM systems and applications. These changes are sparking a symbiosis between user experiences and data-centric governance. On one side, new user-centric experiences are driven by new technology trends, on the other side, more diverse demands for data to drive operational and analytical decisions require higher governance.

Panamerican Mathematical Journal
ISSN: 1064-9735
Vol 34 No. 4 (2024)

The challenges driven by the increased use of data can be addressed by exploring new experiences influencing MDM. These experiences typically focus on crowdsourcing and data utilization. For crowdsourcing, organizations are especially using the strengths of employee interactions, either in the form of tools, gamification, or training but increasingly also acknowledge the quality of customer contributor experiences. These experiences may include direct and easy-to-use ways for customers to create and update their data, enrich product data utilized for personalization, and collaboratively share experiences that enhance the quality of master data, and trust in an organization's commitment to its customers. Data utilization focuses on encouraging the use of quality data by increasing the visibility of data assets and presenting their value. With new technologies such as Digital Twins of the Organization, organizations are spatially connecting business processes to datasets, making it easier for employees to understand what data assets are available for which decisions.

Equation 3 : Master Data Integrity $$\mathcal{I} = \lambda_1 \cdot \mathcal{V} + \lambda_2 \cdot \mathcal{C} + \lambda_3 \cdot \mathcal{T}$$

where

$\mathcal{I}$ = Integrity score

$\mathcal{V}$ = Value consistency

$\mathcal{C}$ = Linkage completeness

$\mathcal{T}$ = Timeliness

λ_i = Domain weights

8. Conclusion

Master Data Management (MDM) systems, like Data Warehousing, Connector, and Quality, enable enterprises to comprehensively manage all parties and things they are doing business with to effectively operationalize important enterprise systems. Knowledge Management, CRM, Product Management, Digital Transformations, Robotic Processes Automation, Big Data, and Predictive Analysis request the effective operationalization of MDM. Intelligent MDM Systems (iMDMS) are at the center of Data-Motivated, Sales-Oriented, and Field-Centric Digital Transformations. They harmonize Customer and Product Master Data (CPMD). With iMDMS enterprises can leverage benefits from intimate Customer and Product Relationships (CPR) and close links between Business Decisions, Sales Performance, and Bottom Line Profit of enterprises.

On the enterprise side, CPMD harmonization reduces Management and Quality Gate Overhead of functionalized planning and execution of business operations. Within the business ecosystem, CPMD harmonization reduces interface complexity and connection costs of each external partner with the enterprise’s Application Landscape. DMS is Data-Driven Middleware between the Business Core System (BCS) and MDM Systems and Key Business Applications (KBA) of CRM, Selling, Order Processing, and Supply Chain Execution. The KBA Management Process (KBAMP) connects BCS Enterprise Master Data (EMD) with globalized, unified, and harmonized CPMD, which are subjected to Client Enterprise Controlled and Computing-Shared Joint Authorities (JAs). iMDMS augments classical MDM, integrating with the iMDMS Management Agent (MA) Error Management Functions of Information Quality (IQ) Root Cause Analysis Procedures. The ”more-less” principle to minimize management overhead of the Smallest Inactive Client Account Management (SICAM) is complemented by the always Secure via Integrated Risk Management of High Impact Client Accounts (HICA) to sustain competitive market-driven differentiation from low-suffering strategic niches up to High Impact Mega Players (HIPL) with Balanced Business Portfolios (BBPs).

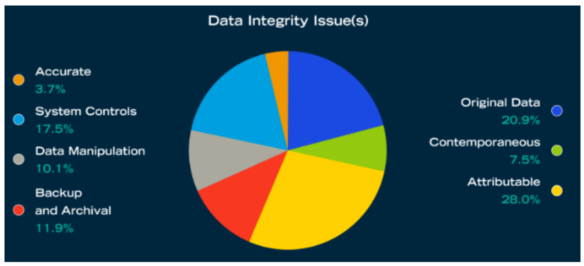

Fig 6 : Data Accuracy with AI-Powered Boomi MDM

8.1. Summary and Key Takeaways

This book presented a holistic, business-oriented approach for moving towards more intelligent systems to manage Product and Customer Master Data in medium and large Enterprises and in doing so provided answers to the questions: Why is PIM/MIM/CDP a Strategic Necessary for Enterprises? How to Enable Intelligent MDM with Service Master Data? Which Core Applications Support a Successful Intelligent PMM/CDP Adoption in Enterprises? This helped to enrich data sciences with a practical and user-centered view from the Market and Sector, from the Core Functionalities of and Key Success Factors for PIM/CDP to Manage Master Data and thus Enable the Customer Experience. These factors, such as the data model or reference content, are the basic prerequisite for the analytics. GDPR-compliant projects, standardized services, and software should increase user acceptance and lead to success sooner. After all the keywords to the answer come together coherently: product data synchronization by best-fit standardization and enrichment plus intelligent shoemaker AI bridging to customer master data, industry, and area reference content. The book is written in a way that citizens and industrial experts can engage for their own needs and understand the complexity of why so few enterprises succeed with their MDM efforts focused on PMM/CDP. It puts both Artificial Intelligence and Content, meaning Trustworthiness and Usefulness of the Additional Data to close the gaps between standard and customer product data encoding harmonized with context-driven Data Governance there. The PIM/CDP Software is a tool; it has to be the Enterprise's Governance for the Data.

9. References

[1] Pandiri, L. (2024). Integrating AI/ML Models for Cross-Domain Insurance Solutions: Auto, Home, and Life. American Journal of Analytics and Artificial Intelligence (ajaai) with ISSN 3067-283X, 2(1).

[2] Koppolu, H. K. R., & Sheelam, G. K. (2024). Machine Learning-Driven Optimization in 6G Telecommunications: The Role of Intelligent Wireless and Semiconductor Innovation. Global Research Development (GRD) ISSN: 2455-5703, 9(12).

[3] Singireddy, J. (2024). Deep Learning Architectures for Automated Fraud Detection in Payroll and Financial Management Services: Towards Safer Small Business Transactions. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 75-85.

[4] Lakkarasu, P. (2024). From Model to Value: Engineering End-to-End AI Systems with Scalable Data Infrastructure and Continuous ML Delivery. European Journal of Analytics and Artificial Intelligence (EJAAI) p-ISSN 3050-9556 en e-ISSN 3050-9564, 2(1).

[5] Inala, R., & Somu, B. (2024). Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).

[6] Inala, Ramesh, and Bharath Somu. "Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making." Journal of Artificial Intelligence and Big Data Disciplines 1, no. 1 (2024).

[7] Yellanki, S. K. (2024). Co-Creation and Connectivity: The Role of Consumers in Digital Ecosystem Evolution. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).

[8] Koppolu, H. K. R., & Sheelam, G. K. (2024). Machine Learning-Driven Optimization in 6G Telecommunications: The Role of Intelligent Wireless and Semiconductor Innovation. Global Research Development (GRD) ISSN: 2455-5703, 9(12).

[9] Lakkarasu, P. (2024). Advancing Explainable AI for AI-Driven Security and Compliance in Financial Transactions. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 86-96.

[10] Botlagunta Preethish Nandan. (2024). Semiconductor Process Innovation: Leveraging Big Data for Real-Time Decision-Making. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 4038–4053. Retrieved from https://eudoxuspress.com/index.php/pub/article/view/2737

[11] Raviteja Meda. (2024). Agentic AI in Multi-Tiered Paint Supply Chains: A Case Study on Efficiency and Responsiveness . Journal of Computational Analysis and Applications (JoCAAA), 33(08), 3994–4015. Retrieved from https://eudoxuspress.com/index.php/pub/article/view/2734

[12] Inala, R., & Somu, B. (2024). Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).

[13] Motamary, S. (2024). Data Engineering Strategies for Scaling AI-Driven OSS/BSS Platforms in Retail Manufacturing. BSS Platforms in Retail Manufacturing(December 10, 2024).

[14] AI-Powered Fraud Detection Systems in Professional and Contractors Insurance Claims. (2024). IJIREEICE, 12(12). https://doi.org/10.17148/ijireeice.2024.121206

[15] Goutham Kumar Sheelam,Hara Krishna Reddy Koppolu. (2024). From Transistors to Intelligence: Semiconductor Architectures Empowering Agentic AI in 5G and Beyond. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 4518–4537. Retrieved from https://www.eudoxuspress.com/index.php/pub/article/view/2861

[16] Mashetty, S., Challa, S. R., ADUSUPALLI, B., Singireddy, J., & Paleti, S. (2024). Intelligent Technologies for Modern Financial Ecosystems: Transforming Housing Finance, Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions. Risk Management, and Advisory Services Through Advanced Analytics and Secure Cloud Solutions (December 12, 2024).

[17] Lakkarasu, P. (2024). Advancing Explainable AI for AI-Driven Security and Compliance in Financial Transactions. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 86-96.

[18] Bharath Somu. (2024). Agentic AI-Enabled Fraud Prevention: Multi-Agent Collaboration Models for Real-Time Threat Detection and Response in Digital Banking. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 4073–4095. Retrieved from https://eudoxuspress.com/index.php/pub/article/view/2740

[19] Yellanki, S. K. (2024). Co-Creation and Connectivity: The Role of Consumers in Digital Ecosystem Evolution. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).

[20] Koppolu, H. K. R., & Sheelam, G. K. (2024). Machine Learning-Driven Optimization in 6G Telecommunications: The Role of Intelligent Wireless and Semiconductor Innovation. Global Research Development (GRD) ISSN: 2455-5703, 9(12).

[21] Gadi, A. L. (2020). Evaluating Cloud Adoption Models in Automotive Manufacturing and Global Distribution Networks. Global Research Development (GRD) ISSN: 2455-5703, 5(12), 171-190.

[22] Goutham Kumar Sheelam, Botlagunta Preethish Nandan. (2022). Integrating AI And Data Engineering For Intelligent Semiconductor Chip Design And Optimization. Migration Letters, 19(S8), 2178–2207. Retrieved from https://migrationletters.com/index.php/ml/article/view/11913

[23] Meda, R. (2024). Predictive Maintenance of Spray Equipment Using Machine Learning in Paint Application Services. European Data Science Journal (EDSJ) p-ISSN 3050-9572 en e-ISSN 3050-9580, 2(1).

[24] Motamary, S. (2024). Architecting Scalable Multi-Cloud Environments for Telecom OSS with Advanced Infrastructure Automation Tools. American Online Journal of Science and Engineering (AOJSE)(ISSN: 3067-1140), 2(1).