International Journal of Finance (IJFIN)
Vol. 36, No. 6, November - December, 2023 pp. 579-606.
ISSN: 1041-2743
website: www.ijfin.com
DOI: https://doi.org/10.5281/zenodo.15804936

Revolutionizing Customer Master Data in Insurance Technology Platforms: An AI and MDM Architecture Perspective

Ramesh Inala
Data Engineer, USA.
ORCID ID: 0009-0009-2933-4411

Abstract

We present a new architecture pattern for the implementation of Master Data Management for Customer Records, as the core of insurance technology platforms. It integrates the initial creation and perpetual trusts of the Customer records, with the native capabilities of Insurance technology platforms. This architecture is contextualized into components, business processes and Digital Twins. We use case studies to illustrate the capabilities of the architecture as a solution to common problems in the industry. We also traverse into automation rules, utilizing Artificial Intelligence technology enabled by neural networks and language transformers. Our architecture aims to enable the automation a deeper customer understanding and segmentation, driven by the detection of the customer voice exposed into the communications and interactions of the customer with the insurance company.

Keywords: Customer Master Data, Insurance Technology, InsurTech, Artificial Intelligence, Machine Learning, Master Data Management, MDM Architecture, Data Governance, Data Integration, Data Quality, Customer 360 View, Policyholder Data, Digital Transformation, Data Analytics, Automation, Real-time Data Processing, Cloud-based MDM, Data Stewardship, Predictive Modeling, Personalization, Regulatory Compliance, Data Orchestration, AI-driven Insights, Identity Resolution.

How to Cite: Ramesh Inala. (2023). Revolutionizing Customer Master Data in Insurance Technology Platforms: An AI and MDM Architecture Perspective. *International Journal of Finance (IJFIN)*, 36(6), 579–606. DOI: https://doi.org/10.5281/zenodo.15804936

1. Introduction

In the present data-centric-oriented market, customers expect that companies provide them strategies, products and services that evolve along with their needs and desires. However, in the majority of cases, insurance companies have huge gaps in their knowledge of their customers and, in relation to its timeline, the damage from Covid-19 has revolutionized our understanding of 'normality' causing a deeper ensemble of economic, social, political and personal changes. For insurance companies, those gaps conduct to poor customer communications, non-efficient marketing performance and trust-breaking user experience. Master Data Management has been used for years to create the single source of truth in companies' reference data. It is responsible for storing and managing data used to support the business transactions. MDM refers to the processes, governance, policies and management of critical data, across the business ecosystem, that affect processes and performance.

2. The Role of AI in Insurance Technology

In the insurance technology sector, the deployment of AI technologies that were made available via cloud apps in SaaS form has achieved rapid gain from early adoption across various underwriting, broking, and claims processes. Initiatives driven by startups with venture capital sponsorships have attracted first customers in these business process segments, and more traditional insurance technology platform vendors with shares of the identifier market or policy administration have either acquired these startups or formed partnerships in vertical markets to add these AI capabilities to their platforms. New investments are now being directed more towards verticalized insurance technology providers focused on developing multi-line, multi-process platforms that address multiple functional areas of the insurance enterprise.

Much of the enterprise effort has focused on enabling and scaling neural network deployments on services that make the machine learning pipeline for building neural networks easier by providing management tooling that reduces the engineering effort required. Marketing, sales, and customer service organizations are increasingly seeking the benefits of building neural networks on customer

interaction data that are trained to supplement or automate human operations across customer journeys. For instance, in customer service, chatbots have been deployed as virtual agents in an effort to respond quicker or offload more inquiries from human agents. Such chatbots may go beyond scripted responses, thanks to the development of transfer learning techniques and proprietary language models, and are broadly termed advanced chatbots.

3. Understanding Master Data Management (MDM)

Master data management (MDM) is the process of managing an organization's critical data to provide, with a high degree of certainty, a single point of reference. This implementation provides business users and IT with the data integration capabilities to support the business processes of a lively business. Most of the enterprises start implementing the MDM with customers because they are the basis for most of an enterprise's transactions. The success of any organization depends on recognizing, understanding, and meeting customer needs.

Fig 1: Master Data Management

MDM can solve some data challenges including inconsistent information about the same element, employees doing the same tasks over and over to prepare bulk loads of data, persons or entities that are duplicate across departments, lack of a corporation's information strategy, poor

understanding of the relationship of chain of customers to their parent corporation, data exchanged between organizations is inconsistent, high cost of data aggregation and analysis when it is located in many different systems, slow and error-prone batch processes that are performed to consolidate data, lack of sufficient data quality tools, so each area is building their private data quality checks for each process that uses operational data, and finally that multiple versions of the same general ledger transaction are being run to create consolidated reports.

Organizations are aware that good data is essential for their market and operational needs. Customers expect service representatives to offer personalized and relevant products, available on-demand, at the most competitive costs, with transactional security. Customer Master Data is central to achieving such business goals. It supports such operational processes as marketing, sales, and service. And it supports business intelligence and data warehousing concerning customer behavior patterns and financial performance metrics.

4. Current Challenges in Customer Data Management

The special relationship between consumers and insurance service providers, shaped through a history of claims paid or not paid, has been often overlooked while building Insurance Technology Platforms. This is mainly on account of the idiosyncratic nature of data related to policies, premium, underwriting, claims, etc., the disparate system technology silos generating this data, urgency for IT system updates without waiting for integration, and the sheer volume of transactions involved. While regulators will mandate storage of customer data for 5-10 years in a compliant manner, the lack of fluency around Customer Master Data Management on the part of insurance service providers has paved the way for third-party claims aggregators to exploit this relationship by using sales and marketing strategies, further taking away business from these companies. This is compounded by the increase in data protection regulations.

Hence, it is very important that Insurance Technology service providers replicate the Claims Customer Experience on their IT systems by enabling the firm to visualize the customer relationship via multi-channel, multi-dimensional attributes, like a customer dossier, Customer 360 view, etc. This would enable these firms to drive personalized cross-sell and up-sell strategies. However, technology replications for a Customer Master Data management are fraught with challenges; not the least of which is the fact that Customer Master Data is typically stored in respective IT system silos. The Customer Master Data may well be stored in Claims systems, Back Office Finance systems, Policy Underwriting systems, Policy Valuation systems, etc. The Customer Master Data is

prone to data redundancy and inconsistency due to periodic updates. Quality data management, physical data integration, and virtual data integration are the various processes recommended to gain access to the information, albeit with relative pros and cons.

5. AI-Driven Solutions for Data Quality Improvement

Insurers have been dealing with data quality issues for many years. Typically, insurers perform small data quality assessments on key areas, such as customer name and address, then create and manage data profiling rules through periodic reports and associated exception handling. A combination of workflows, logic models, and reporting dashboards can be used to monitor these exceptions and help resolve data quality issues. These solutions are generally IT-driven, don't scale, and are based on business rules that are not easy to create and manage. More importantly, these solutions don't eliminate the issues that are noted, as the business is not actively engaged.

Further, with the advent of online mobile transactions, time is of the essence, and these small, manual, periodic solutions delay transactions. A new, automated approach is required that uses an enterprise-wide analytics engine to identify, correct, and respond to intraday data quality issues. The solution should be flexible enough to perform predictive data quality problem moderation and apply models to prevent poor data quality at the source. It would have an easy-to-use interface for business users, Data Stewards, or Data Quality Analysts to identify data quality issues easily. It has to be comprehensive enough to recognize types of issues that are currently not covered by other solutions. It must also differentiate between business rule violations and factual errors from trusted sources, including internal applications or external trading partners. Most importantly, it should support a continuous improvement process that leverages the business to tune the rules continuously, enhancing user experience.

6. Integrating AI with MDM Frameworks

The proper functioning of any insurance organization depends primarily on the quality of its customer data. However, in today's fast evolving technology landscape accessing customer information via traditional frameworks is highly limiting if not altogether impossible for companies today. AI technologies today have made it possible for companies to properly catalogue, update, curate, enrich and even discover information about their customers. We propose the Fusion

Framework which couples together AI technologies with a traditional framework, and allows for a multitude of activities aimed at revolutionizing Customer data.

As organizations harness customer relationship management solutions to gain insight into customer needs, and deploy websites, call centers, and self-service kiosks to enhance customer interaction, there's a lot of customer activity data being generated. It is critical that the organizations develop well-defined repository infrastructure for customer data at the center of all operational processes, so that data is accurate, up-to-date and has a consistent single source of truth. However, as the volume and variety of customer transactions increase, maintaining data as trusted and verified becomes an uphill battle. In this report, we present a paradigm for revolutionizing Customer Master Data Management solutions. We utilize a variety of AI technologies and information functions in the Customer space to enable a self-healing and self-curating repository of customer data. Examples of Information Functions are recommendation engines, learning to rank, active learning, probabilistic data association, record linkage, and attribute matching. By using these Information Functions, we build a system that is used by operational systems and these same Information Functions are also invoked in the process to help cleanse, reconcile, and enrich the data as it is being ingested. Data entry is an important part of the process and it could be a source of a lot of errors, but with a self-improving recommendation system, we could make it almost seamless for the users and remove the pain of bad user experience.

Equation 1: Master Data Entity Resolution Score

$$\mathrm{ER}_{\mathrm{score}}(i,j) = \sum_{k=1}^{n} w_k \cdot \mathrm{sim}(x_{ik}, x_{jk})$$

- x_{ik}, x_{jk}: Attribute k of customer records i and j (e.g., name, address, DOB)
- sim: Similarity function (Jaro-Winkler, cosine similarity, etc.)
- w_k: AI-learned weight for attribute k based on predictive power

7. Data Governance in Insurance Platforms

Having an accurate view of your data is critical to upfront and continuous process excellence within insurance underwriting, claims management, and customer service. Having that view requires data sharing and trust across departments. Data governance is an established tactic to allow

departments to collaborate more fully in sharing responsibility for maintaining a reliable and accurate repository of data. Given the scale of investment and the centrality of data to insurance value creation, data governance efforts will be imperative. Data governance is an effort best placed within the business environment. Enabling governance is most likely best supported by a central office in IT that chartered the governance frameworks, has oversight for adherence and review of exceptions, supports data management and quality operations, and provides and governs central data repositories for the lines of business.

Governance should extend not only to such specialized collaborative functions but also branch out into data definitions, quality metrics, resolution of data conflicts as they arise in business, systems, and workflows. Organizations should have defined change control in terms of the department or corporate resources that fund development. Those resources should have a collaborative role in reviewing, testing, and providing sign-off to deployments of platform releases that involve major feature changes-characteristics affecting data, process, service, experience, or system of customer record in MDM – as these features address business issues with corporate visibility, such as data quality improvements through data augmentation or cleansing and de-duplication, enhanced workflow UX design, improvements addressing performance or availability issues, and enhancements to datasets such as devices and experiences.

8. Case Studies of Successful Implementations

The insurance industry's complexity and volatility present challenges for implementing architecture changes in insurance technology platforms, especially in converting siloed applications to unified platforms. This chapter highlights examples of insurance technology platforms re-imagined with innovative solutions that simplified the customer and other master data definitions and functions. New approaches included refining the implementation of an insurance technology platform as an insurance data fabric in a cloud-native ecosystem architecture and moving the enterprise data warehouse to the cloud, with a federated data structure and access layer to consume the customer master data and optimize customer service.

These insurance technology platforms integrate the core icons and use point solutions to add functionalities. A specific implementation enhances the customer journey with central access for insurance agents, underwriters, customer service representatives, and customers to a federated data layer integrated with customer master data and industry reference data that aids insurance

operations. Data quality processes and validation judgements, enhanced by artificial intelligence, support the decision-making process for functions like customer onboarding and service requests.

This integration improves the customer experience by minimizing data sharing, thereby reducing friction in the development of customer journeys, and supporting technologies like bots and robotic process automation. The solutions can break down barriers between functions in insurance operations like sales, underwriting, claims, and customer service. Customer master data and its use cases are key pillars of these implementations, which enhance customer service and make the customer experience seamless.

9. Technological Innovations in MDM

The conventional MDM tools and technology stacks, based on the 'one-size-fits-all' monolithic application approach, have proved to be limited in their ability to meet the demands of the digital business era. Newer versions of existing tools, with varying degrees of sophistication in their capabilities, are popping up in the MDM technology landscape. Moreover, mature tools now provide MDM capabilities as features or subset functionalities of larger enterprise systems, such as Cloud-based ERP systems with data management capabilities, customer experience management tools, and big data ecosystem tools. Lack of enterprise data strategy coupled with the high costs or risks of deploying an MDM application has resulted in data-driven enterprises deferring the MDM technology decision for too long. Forcing the decision later has made the task more complicated. With the emergence of significant diversity in enterprise process and customer interactions, data management needs are increasingly specific to use cases driven by industry-specific customer interactions, types of data involved, and specific customer engagement channel. The bottom of the pyramid, or smaller digital disruptor companies entering customer-centric business processes, generally have dissimilar unique needs in deploying customer data management. Their decisions are influenced less by considerations of cost but more by the need to manage speedier deployments to enable differentiation.

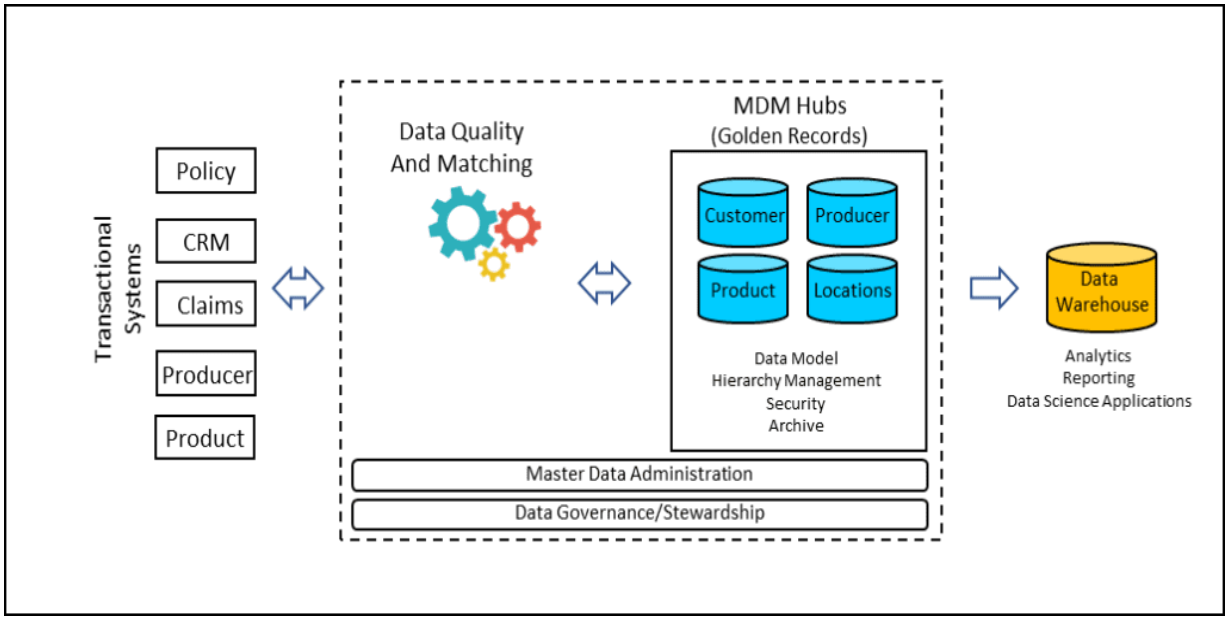

Fig 2: Achieveing Clarity in Customer Data Management for Insurance Industry

Insurers want customer master data to be a strategic layer in their enterprise technology stack forming the foundation for flexible business process and insight enablement. This market demand is challenging software and technology suppliers to provide MDM technology architectures that balance the need for enterprise infrastructure robustness with that for data management capabilities specific to unique use case business process needs. In the domain of enterprise cloud software, a layered architecture has already enabled a lot of the requirements for specific use case-driven enterprise customer master data management. It provides a generic infrastructure for enterprise customer engagement built on its cloud application platform, common industry use case-driven applications, and common industry-specific templates, on which third-party partners can build vertical applications for industry-specific engagement.

10. Impact of Regulatory Requirements on Data Management

Work in Progress - Please see personal communication for more context and information Regulatory requirements can impact many dimensions of data management - from where and how data is stored to who gets access, and for what purposes. Some regulatory requirements have been around for a longer time and have been incorporated more naturally and organically into the data management strategies of insurance companies. Examples of these requirements are KYC, or Know Your Customer, requirements that arose from banks' and financial services companies' need to have

holistic views of their customers' transactions and activities, and privacy-related requirements, including data restriction and deletion requirements. Other regulations arise from more recent events - like the sudden shift to need for claims during the pandemic, and the increased risk associated with natural catastrophe events. These regulatory requirements usually relate to the speed, flexibility, and the nature of services offered by insurers. While Insurers have been somewhat insulated from the impact of these events in the past, the rise in ecosystems-based services in insurance are likely to make them more susceptible to disruptive change.

Access to data is essential to forming a holistic, 360-degree view of customers by insurers, brokers, and other intermediaries within an insurer or insurance ecosystem. However, privacy, conflict of interest, as well as many other reasons restrict access to all available data. Data processing and commissions boards, as well other trusted third parties, usually initial role in servicing requests for data from various participants. Each new regulation creates new guards and rules to control how data can be shared in the name of safeguarding security and privacy – but these rules without an equally robust approach to automating scrubbing sensitive information can act as a further barrier to enhancing and uplifting data within a firm’s Datalake.

11. Customer Experience Enhancement through Data Management

Master data management and the notion of the customer master data in insurance technology depend on data quality. Despite the fact that various insurance partners now have access to a large volume of data on a single customer, including data on their personal characteristics, important life events, and unique behaviors; insurance companies that rely on consumer trust have been appalled by the adversarial nature of data erosion and incorrect recommendations. For insurance technology companies, effective usage of master data remains a key issue. Customer-by-customer master reference data accuracy is the basis of enterprise information initiative effectiveness. A high prevalence of master reference inaccuracies correlates positively with poor sales initiatives; incorrect service recommendations; weak risk evaluation indicators; and other perilous issues for modern enterprises.

Enterprise data management technology must assist in solving these insurance technology business challenges – an impossible aspire if advanced capabilities are lacking. Yet many companies still face internal obstacles that hinder progress. In fact, while a significant percentage of executives think that good data management yields competitive benefits, nearly the same percentage cannot execute on a data strategy. What causes this disconnect? One issue comes from the complexity of

managing data quality on a global scale within large, impermanent ecosystems. Another issue arises as corporate plans evolve rapidly, leaving less governance. Without centralized control, it is hard to understand and improve data that cross organizational boundaries. Meanwhile, because many business interactions are now handled in shifts, data needs frequent updating. But most executives concede that they do not have the tools they need to collaborate effectively.

12. Future Trends in AI and MDM for Insurance

With the rapid technological advancements, especially in the cloud and data and computing space, AI will also have fast changing and improving tools and methods to evolve both Master Data Management capabilities, and customer data management tasks and processes. Cloud-based MDM will enable more frequent and on-demand processes to be setup and configured, rather than inflexible and one-time only implementations, such that even processes that have data at the core, can be quickly created based on ever changing business needs, and be more agile. This applies to customer data in insurance as well and will lead to more automation and machine learning applied to customer processes in insurance. All of these will address the part of the AI promise that talks about less human effort in doing such repetitive tasks. Less human effort on such automation or human-in-the-loop work does not imply relaxing quality. As such technology is democratized, but the cost of errors is not compromised, it will more and more lean on such automation to drive customer experience. With MDM as a platform that can support the training, creation, and optimization of the rules enshrined in the AI, this leads to a symbiotic relationship between AI, MDM and data in the MDM, especially for customer data.

At the other end of the spectrum, to make insurance companies more innovative, we think that a deeper and more effective engagement with customers would help insurance companies capitalize on opportunities such as protecting the customer and increasing the share of wallet in integrated customer engagement. Such an approach for insurance companies would also help address the issue of long-term sustainability for insurance companies. This brings us to the converse trend in using AI and MDM together for insurance – one with more human effort and integration, especially to execute the engagement strategy. The unified view of customer across channels, segmented, analyzed and prepped for customer interactions is key to this strategy.

Fig 3: MDM Essentials Across, Process Ans Technology

13. Data Privacy and Ethical Considerations

The amplified functional, operational, data management, and technical demands placed on customer master data assets, combined with increasing privacy awareness and legislative regulation of consumer data, has created a polarizing standpoint for mutual benefit. Today, consumers remain eager to rapidly engage in their demanded personalized experiences, but are concomitantly wary of the necessary data sharing required. They expect the benefits of insightful personalized relationships and engagement across business verticals, backed by increased transference of data wealth. On the opposing, regulatory implementing bodies, along with its judicial review and enforcement capabilities, pursue entity required commitment to scotched physical, data, and emotional engagement crossover of non-sensitive data deserves. Whereas enterprises strive to augment the value and service they supply downlayers in the microsocietal consumerism ecosystem, concise and well-structured governance policies on consumer data privacy that encumber their platforms, especially customer master data, serve as enabling ordinance.

Against the consumerism backdrop, it is vitally important for the insurance enterprise to proffer consumer-centric, trust-building commitment to data stewardship. This balance can only become actualized via robust master data. Consumer-centric privacy regulations set the bar for insurance organization deployment of technology-driven, data management enabled, transparent policies addressing the ethical, immoral, and illicit collecting and use of not only customer identifying information, but their transactional information, behavior knowledge, and consumer

sentiment relativity. The industry accepted ethical imperatives are simple to espouse, yet require indomitable corporate governance to achieve carrying out. Organizations must tightly control datasets to protect them from illicit access and use inside and outside their ecosystems, and create strict protocols surrounding informed and empowering consent.

14. Building a Robust MDM Architecture

Robust Master Data Management (MDM) architecture is the required backbone of every enterprise technology platform. It helps ensure that all other services of a platform are using the most accurate and reliable trusted data. The core guiding principle of a good MDM architecture is to expose APIs which are easy to use and remember for the other platform services. These services should contact the MDM APIs to seek answers related to trusted data. The MDM architecture can be designed for the customer domain model used by the enterprise or the policy object domain model. The customer domain model is used widely during the application development since every enterprise provides services to a customer, whether it is the agent or end policyholder. Policy object domain model is primarily used during mapping of the customer data provided by the partner system with what is already available in MDM. Therefore, MDM architecture must also provide the ability to map master data between multiple domains.

MDM can be designed in a decoupled architecture where the MDM service is treated as an independent service externally but uses a workflow orchestration engine to flexibly connect to the transactional systems of different business units and contact them as necessary to collect the data, validate the data, and execute policies to store them in a database of choice or a data lake as necessary. The reliable persistent master data created and curated in these databases or data lakes must then be exposed as APIs to the other services who may be written in different languages, and so these APIs must follow standards as it exposes more realistic query and data filter capabilities.

14.1. Key Components of MDM Architecture

MDM architecture takes into account solutions for all aspects of the data lifecycle. Such solutions usually contain components to create, upgrade, and delete datasets and the relations between these datasets, workflows to govern MDM operations, custom modules to adapt MDM components and workflows to idiosyncratic use cases, connectors to push/pull data to/from MDM systems and other IT systems, a front-end interface for data stewards to manage data and operations visually, tools to obfuscate/manage personal data, and tools to audit changes to MDM datasets and

tracks MDM operations because both auditing features are essential for privacy management in general and for operation control in highly regulated industries, such as insurance, in particular. In addition to the above components, a production-ready MDM architecture needs to consider Power and Cost management for customers who save massive amounts of money by hosting cloud servers on the least busy days, High Availability to accomplish the "always on" requirements posed by customer operations, Backup and Disaster Recovery to offer a reasonable RPO and RTO service, Performance and Monitor to make latency of queries and other functions as predictably short as possible, there are checks that trigger alerts if one aspect is not as close as expected to defined thresholds, such as the number of logs, Operations Efficiency, DevOps and Continuous Integration to shorten time-to-market for new developments. Because of the rich dependency on other IT systems, this is facilitated if MDM cloud components can easily and reliably exchange data with external data sources and sinks of data used in MDM functions from other IT systems.

Equation 2: AI-Driven Data Quality Improvement Rate

$$DQ_t = DQ_0 + \sum_{i=1}^{t} \Delta DQ_i \quad \text{where} \quad \Delta DQ_i = \phi \cdot \text{ML}_i$$

- DQ_t: Data quality at time t
- ϕ: AI effectiveness factor
- ML_i: Model-driven correction at time i

14.2. Scalability and Flexibility in Design

A Cloud-enabled and refactorized Microservices based design built for ‘Data as the core of Insurance Technology Platforms’ to Power AI-enabled Next Gen Insurers is necessary to ensure that Master Data Management Platform can be kept in Sync with Supply Chain Management for Insurance Business Capability Models. Enabling a scalable architecture is very crucial as Master Data acts like the oil or fuel to the engine of Insurance Technology platforms and if it is not in its purest form, it will damage the entire system of insurance technology models. Presenting the Master Data in the most simple and easiest way to be used within the technology ecosystem of Insurance enterprise enabled by ease of modulation and Modeling is critical for the success of any organization and its master data management initiative projects.

MDM Software should be flexible enough not only to model the data based on business capability needs for posing new business capabilities to marketplace, but also bring in flexibility to

provide Custom Modules at lightning speed to Business capability in terms of performing new business, new opportunities, new services, New Revenue operations and bring in all channels of business capability like portfolio management supported by the external partners situated both inshore and offshore whether be it B2B or B2C, to ensure that MDM acts as the fuel to the business functionality and perspective instead of becoming a software service that controls the business. Further, the data elements that are important for insurance underwriters and actuaries should be easily accessible and consumable directly from the MDM Solution through APIs to their Data Science Model management, complexes and Maturity Models.

15. AI Algorithms for Data Matching and Deduplication

Data quality is fundamental for Generative and Foundation models, particularly when these models handle dangerous domains such as financial systems and solutions, medical applications, and others. Insurance core systems, and in particular Agency and Customer Relationship Management solutions, many times are affected by large data quality issues. For example, millions of dollars are spent each year trying to recover from fraudulent claims performed by criminals that purposely create their identity in the system using a different address or a small variation of a name. This bad actor behavior creates Insurance Business Entities with overlapping details: aliases, names of relatives, names of companies or LLCs, addresses, geolocations, phones, emails, and others, engaged in different insurance activities. Investigating this activity can be a long, expensive, hard, and dangerous task for insurers that own the claim. Moreover, during normal business, insurance companies are exchanging and sharing information about companies and individuals in order to detect M&A related fraud as well as collaborate to avoid fraudulent activities.

Investing in Identity Technology is, therefore, a key priority for insurance companies. Foundational Insurance as well as Insurtech companies are making heavy investments in advanced AI models and proposing services for customer identity management. With this goal in mind, an ideal ACRM system would contain advanced address verification with Natural Language Processing and Soundex algorithms; deep deduplication tools, usually based on supervised Machine Learning, specialized embeddings and later filtering of records that are considered near duplicates by traditional algorithms; enrichment of customer profiles with under-the-radar data that would allow creating true customer records even without SSN; and finally true anonymized IDs such as Knowledge Graphs that allow tracking customer movements across different IDs and across time.

16. Real-Time Data Processing in Insurance Platforms

Many industries rely on batch-based ETL processes for collecting, staging, ingesting, processing, and storing their operational data but adopting this approach can make less sense for insurance firms. Long-established commercial data platforms for Policy Administration, Billing, Exposure Management, Claims, etc. typically deal with business events that have been turning into past-tense data before evening batch windows. And while periodic latency of not just hours but usually days can be accepted for revaluating existing closed claims, generating estimates of IoT-connected homes for underwriting dynamic premium credits and detecting what is today a potential tomorrow's ransomware attack targeted at banks requires different data architectures. Earlier real-time EAI solutions from the nineties were invented to orchestrate and automate the mission-critical customer-facing business operations of enterprises so that those operations could be more efficient and less error-prone. They focused on the process how all the business applications have to work together, not on enriching customer or product or claim implausibly semi-automated ETL pipelines running outside the business applications.

Domain master data for insurance represents a shared, distinct, curated, highly structured global view of customers, products, claim events and the relationships between different key entities, often enriched by external data vendors and risk scoring systems, and facilitated by an MDM system. The ever-changing object-centric event data created by insurance companies' various transaction-processing application platforms is not the MDM system's "source of truth" but the first multi-gestalt model to accurately represent important temporal and spatial characteristics of the data. Enterprise integration patterns, such as outbox table integration patterns, event sourcing and CQRS, asynchronous message brokers, hybrid ETL technologies, as well as enterprise application platforms for orchestrating and managing the overall real-time event data supply chain between MDMs and iPaaS connectors, will shift accordingly by serving the widely distributed and extremely heterogeneous application architecture landscape of insurers.

17. Challenges in Implementing AI Solutions

Despite the various advantages of AI technologies, there are also certain challenges of implementing the AI solutions. There are myriad doubts of organizations regarding AI, such as will AI be able to replicate the human judgment with regard to certain critical tasks. Organizations are not only concerned with the quality of certain outputs generated using the AI but are also concerned regarding the security issues and cost that are associated with the implementation of AI. In general,

the challenges in implementing AI technology are as enumerated below:Security: AI is vulnerable in certain aspects, such as the provision of data to third-party companies and the possibility of hacking. Security is one concern that organizations need to address while implementing AI.Cost: The implementation of AI projects is very costly, which may dissuade organizations from implementing AI projects. However, the long-term return on investment made in AI technology offsets the high costs usually associated with AI implementation.The ability of AI tools to produce outputs as per the standards laid down by human beings is still a matter of doubt.

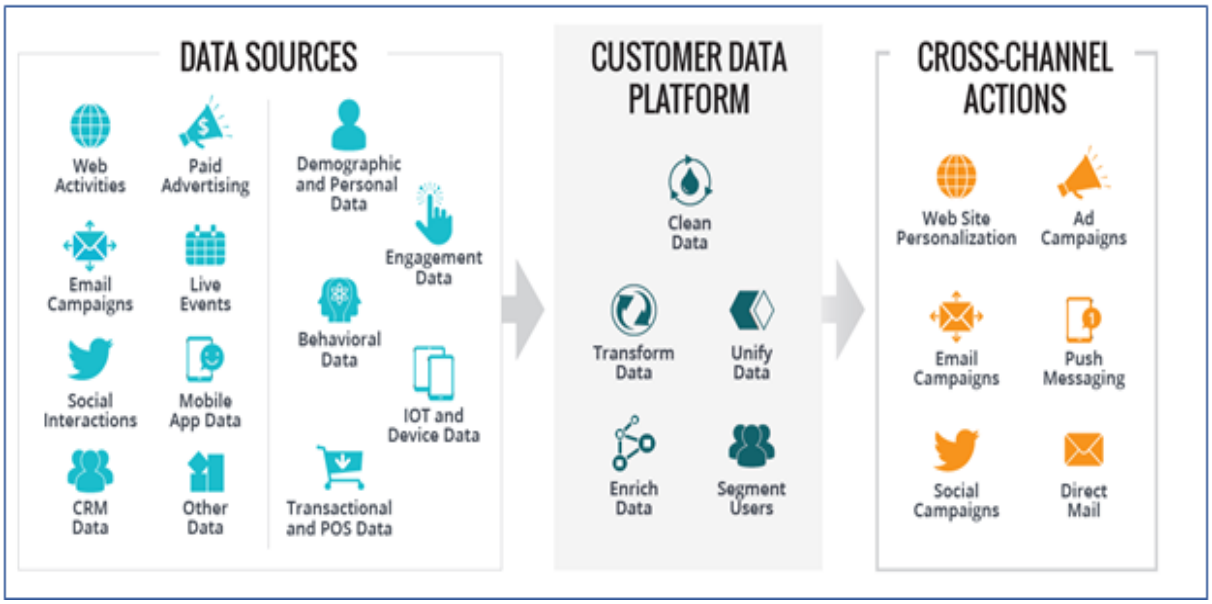

Fig 4: Customer Data Management for Insurance Industry

Energy Consumption: The energy consumption of the atmosphere might increase as a result of the increased demand for electricity from the cloud services and data centers that are essential for AI technology. Companies and organizations must ensure that energy consumption can be curbed while implementing AI.

Bias: An AI model is as good as the data it has been fed when being trained. If any bias is present in the data used for training, the AI model being generated using that data will also be biased. Organizations need to proactively remove the bias that is present in their data before they implement AI projects. If organizations do not remove the bias present in their data, problems may arise while analyzing the outputs generated by the AI models.

Transparency: Another major concern regarding AI is regarding its transparency. Organizations need to take elaborate steps regarding creating an ecosystem of explainable AI that

sends out clear messages regarding the objectives, methods, outcomes, and performance assessments associated with the AI systems.

18. Training and Change Management for Staff

For the Master Data Management (MDM) system to be able to impact the business processes with improved quality Artificial Intelligence (AI) algorithms using clean and normalized customer data, it is crucial that the staff working on the MDM solution is trained on what MDM is and what value it gives to the business as a whole along with AI powered applications so that the processes using the enabled feature can be effective and efficient enough to provide that output. The MDM vendor should create a special training program for Business Process Owners around what data is important to the application and Business Process Logic with rules that should be tied to the master data for cleansed and validated data to flow into the application and the transactions in the applications and to be further processed and how it can affect the AI training models. This training should have special focus on why the MDM solution exists and why the data in the source systems servicing the application should be validated and cleansed with the rules.

To enforce the business process owners to take interest in the process of training the models with clean data, KPIs related to the AI algorithm models should be tied to their Business Process area. The organizations using these associated AI applications should also focus on creating a Center of Excellence (COE) which should contain the skilled resources who are well trained on how MDM solutions exist in the ecosystem and work with the associated applications. This COE should work along with the Business Process Owners Office and have regular communications to check the health of the models and reinstate the importance of clean data and cleansing the data in source systems and how these have an impact on the AI Model.

19. Cost-Benefit Analysis of MDM Projects

The cost-benefit analysis of MDM projects must be at least as rigorous as that of any major capital project. Although it is tempting to justify an MDM investment solely in terms of savings, this would be misleading. Consider the new MDM system savings because of increased productivity across the enterprise alone. The array of systems that probably involves contactable customers, including but not limited to customer service systems, underwriting, policy administration systems, premium collections, claims negotiations, marketing systems, and even agents, will take increased

loads for an enterprise to justify its MDM investment. The increased efficiency that results from a customer-centric (rather than product-centered) view must be analyzed and then justified. Furthermore, the benefits of customer analysis projects are not easily quantified, nor are they specific to MDM.

A benefits model that takes just this consideration into account views an MDM investment as a generic enabler of customer information, thus paving the way for a variety of customer-centric applications. The benefits of specific applications can then be used as inputs to project evaluation. The business cases for customer-centric initiatives, such as relationship management, can be quantified to make the business case for MDM. In turn, the business cases for industry-specific MDM applications can be made more robust by relying on the flexibility of MDM systems. Generalizing these recommendations into general principles, we propose a three-pronged approach to MDM cost-benefit analysis.

The first step in addressing the costs involved in initiating a new MDM project is to consider the enterprise expenses linked to installing existing product-centric databases. Besides the costs of IT to actually implement and maintain the systems, the wider enterprise costs come from the impact of the systems themselves. The benefits of a centralized, standardized MDM implementation need to compare favorably against those incurred from managing disparate, embedded product databases.

20. Collaboration Between IT and Business Units

Due to new technological desires and capabilities, it may be time to revisit some ancient practices surrounding the IT function. Conceiving the business in a decentralized way - and the IT function as a valued trading partner - enables the business to empirically validate ‘pocket-decision-making’ relative to the project on behalf of the entire collaborative team: business + IT. In this manner, the business can dispose of strategic initiatives for which it has the responsibility/authority. The vision/goal must be defined in joint business/IT collaboration. The plan to clear the path toward the ultimate goal must be defined in joint business/IT collaboration. The ultimate technical solutions must be defined by IT in collaboration with the business. The business must ROI the solution - that is, it must have a team that implements, invests, and then crucially and riskily, actually utilizes the new technology. Perhaps the distributed nature of determining what solutions to pursue could be accomplished by developing ‘knowledge centers’ on a functional basis. Instead, ‘knowledge centers’ would exist not on a PM basis, but on an overall business unit basis. Established procedural mechanisms would exist which ensured that IT was always up-to-speed on what was possible within

the currently 'business knowledge-center'. Through this process, opportunities to implement technologies on a project-by-project basis would arise.

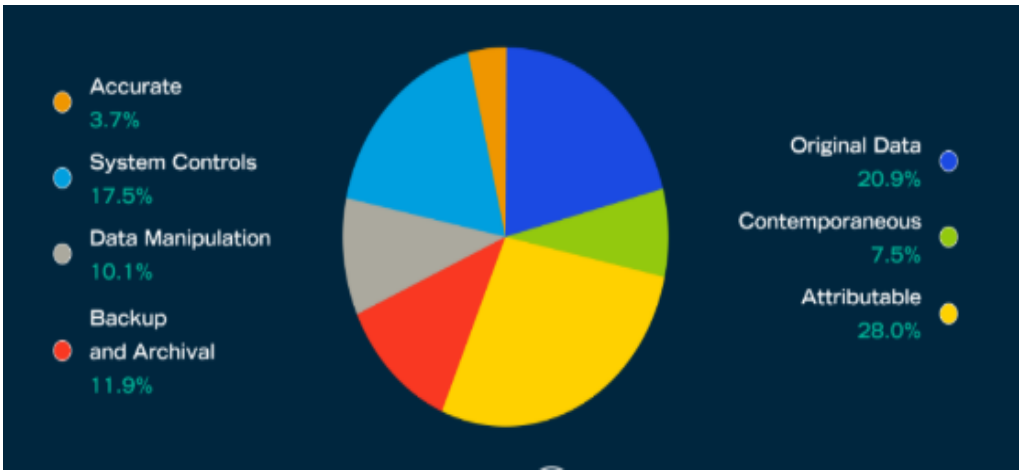

Fig 5: Enhance Data Accuracy with AI

It is our contention that consultation could provide the enablement necessary to develop the group-mode knowledge across multiple knowledge-center domains so that initiatives could be developed to both digital concept and a built system. Using 'virtual' resources in bridging the capacity gap on a knowledge-center level allows quicker bootstrap and digital entrepreneurial development of the core business. We believe that some combination of consulting-based programs linking business units with tech firms must be established that would allow a protective partnering form of entrepreneurial bootstrap.

21. Vendor Selection for MDM Solutions

This section provides industry insights through use cases on considerations for selecting MDM solutions supporting transformation initiatives in the insurance industry.

MDM Solutions are complex systems integrating multiple technologies to address a specific business capability such as customer onboarding, product pricing etc. These MDM use cases touch upon several enterprise systems. The MDM vendor choice must balance between out-of-box connectors, in-house custom coding capabilities, and pre-defined business process models offered by MDM vendors. The vendor must also support the insurance domain specific scenarios, and be able to integrate seamlessly with the internal systems. Several popular MDM vendors focus on defined use cases such as customer onboarding insured vehicles at agent level, customer onboarding

insured property at agent level, benefiting from exchanges, anticipating cross-sell & up-sell products, data curator for customer enrichment & deduplication, etc. MDM solutions can be broadly categorized in the following segments: customer onboarding insurance, onboarding complex services, anticipating cross-sell product requests, customer enrichment & deduplication, etc. In this section, we will evaluate solutions from vendors in these categories and categorize them accordingly.

Selecting the right tools for building the MDM infrastructure is critical. Some MDM requirements may need specialized tools from different vendors, while others are supported through a more generalized but less specialized, more monolithic set of tools. In practical implementations of MDM infrastructure in insurance core technology systems such as policy administration, claims, or underwriting processing, many of the solutions selected for master customer & account data have to work collaboratively utilizing APIs and enterprise service buses.

MDM is a multi-faceted activity and as such has tools and solutions from several vendors. Further some of the MDM capabilities for fulfillment, MDM data master-data-enrichment, MDM data curation, MDM customer-anticipatory cross-sell, etc. may use a glass-box, but a heavy glass-box, where Glass-Box is used in AI-platforms making available the internal details of how the solution works, to insurance technology architecture developers while developing the workflow for that MDM capability.

22. Performance Metrics for MDM Systems

To build on the metrics presented in the previous section, we will be using many of the attributes that are used in the domain of information systems. For example, quality of service (QoS) are often used as performance measures for assessing the effectiveness of MDM services. The operational dimensions relate to user queries and often depend on the parameters of the query such as the size of the data sorted or filtered queried. For example, response time, throughput, up-time, scaling-time, and recovery time from failure. Strategic metrics relate to the organizational impact of providing MDM services and hence are not easily computed. Finally, the administrative measures relate to the internal operations that are used to both tune the MDM service and also reduce the operational costs incurred by providing the services.

In proposing the MDM architecture for business data reconciliation in extended enterprise, a set of performance metrics that build on these dimensions has also been proposed. For example,

information architecture metrics, such as the number of business entities that need to be reconciled for an enterprise, need to be exported from the central MDM for access from the subsidiary systems. Informatics metrics, such as data synchronism, system throughput and consistency of imported data, also exist. Informatics looks for attributes such as consistency, accuracy, completeness, integrity, and redundancy. Variations of some of the metrics such as the time required to create the MDM record in which the instances of the business entity co-exist need to be defined in terms of specific business events. Some of them, such as data synchronism are more relevant to the operational aspect of the MDM service, while others, such as the storage size of the consolidated MDM record impact the strategic aspect.

23. User Adoption Strategies for New Technologies

User adoption is often related to a change in the way employees are accustomed to doing their jobs. Adoption issues are recognized as major factors in the success or failure of new technology projects. Many technology projects fail to meet their objectives; others are abandoned altogether; and many of those that are accomplished do not produce the anticipated improvements in performance, productivity, or profit. While solution design, service engineering, cost management and productization play a big role in ensuring success in technology, it is increasingly clear that organizational factors are vital too. Firstly, during the project, many technology factors come together with organization design. Projects will impact on the way projects have to run, on the balance between technology and humans, and on the deployment process. Key clusters of issues are team working, human talents for the project and organization for the project itself, learning, innovation and creativity, communication, changes in structures and frameworks, and the introduction of new business processes.

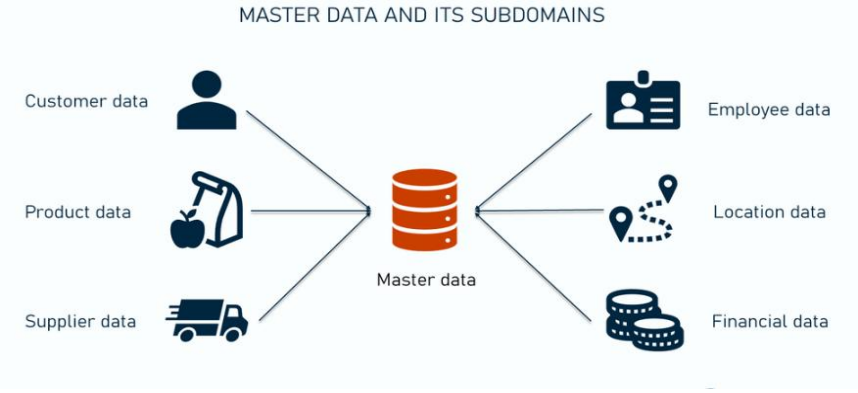

Fig 5: Understand master data management

User adoption issues, policies and actions involve the whole life cycle of a project or change. They need to be considered upstream of the project when developing the business case to support the investment. Therefore, customer, enterprise and project managers need facts, assessments, opinions, experiences, and insights from adoption experts to complete strategy, project and planning formulations initiated by technical contributors who design the investment or project recommendations. During the project life cycle, the technology team must share changes with business line leaders and prepare them for user adoption.

24. Continuous Improvement in Data Management Practices

Achieving and maintaining high data quality is too important an effort to be ignored especially when such emphasis on data management principles are conveyed from top down. Communicating the value of high-quality data and managing services around that for the sake of internal customers or business need becomes important for long term success. Organizations must utilize the best technology and seasoned implementation strategies from reliable partners and consultants while implementing a MD program. Utilization of the best technology must also include other best practices and principles while dealing with data and issues within them. Continuous improvement must be promoted as an organizational value and not a mere slogan. Implementing data governance within the organization becomes important. Organizations continuously examine the MD systems to assess both technology and program goals performance. Technology addresses the systems housekeeping and asks questions like Are we looking to tie into MDM as part of company or merger? Are we looking to integrate MDM into ERP or Business Intelligence? Is someone ill prepared for data requests? Is the MDM project a management failure? Business Goal Performance asks Are we utilizing data more intelligently today than we did yesterday? Are sales taking more advantage of data today than yesterday? Are we producing more upsells today than we did yesterday

Equation 3: AI-Based Customer Segment Assignment

$$S_i = \arg\max_k \; P(Y = k \mid X_i)$$

- S_i: Segment assigned to customer i
- X_i: Feature vector for customer i (demographics, policy history, claims, etc.)
- Y: Segment label (e.g., High Risk, Early Retiree, Corporate Group Member)

The long term philosophy of continuous improvement is reflected in the empirical intent of seven quality management principles and core elements. A concise, business focused explanation of the seven principles is available. Any organization should utilize business critical guidelines to help determine the priority of data management efforts within the enterprise and also on their continuous improvement cycle of data management initiatives.

25. Evaluating the ROI of AI and MDM Investments

The cost to implement a comprehensive enterprise-scale AI and MDM architecture solution can be easily gauged based on estimates or proposals. The real challenge is to understand the benefits. What are the tangible and intangible benefits that an enterprise can expect from such an implementation? Clearly, the follow-on question is, what is the return-on-investment (ROI) of such tangible and intangible benefits? As you can imagine, the answer is not trivial.

There is no doubt that with the arrival of NoSQL databases, there are alternatives to the recent previous so-called big-data-stacks. However, in reality, the AI value-chain companies invest millions of dollars using an AI-based data pipeline architecture which actually does the aggregation and curation to build and manage the AI value-chain. Monetization involves selling to companies from diverse industries and business domains. Monetization markets use different lengths of data; data length is important as it dictates the MDM capabilities of resolving such data with and across companies in different AI value-chains.

Traditional methods of estimating ROI try to measure either realized value directly or economic value indirectly. When trying to identify the realized value, there are typically many challenges, including the identification of the associated revenue streams as well as isolation of the impact of the investment over and above other concurrent factors affecting revenues.

26. Conclusion

The global insurance industry is undergoing a massive transformation, but a major obstacle in the path of Insurance Technology platforms is the disparity and growing complexity of Customer Master Data in existing industry architectures. Customer Master Data is not only growing exponentially but is also becoming heterogeneous distributed and dynamic. Insurers are using disparate systems for internal interactions, and external interactions – marketing, distribution, underwriting, and servicing; while relying on third-party firms for fraud detection, identity

validation, risk and loss assessment, and technology support. With the advent of digital wallets, the sources of Customer Master Data have extended beyond policy accounts to a customer's digital infrastructure. Consequently, customer relations will be built on Digital Customer Master Data.

An alternate solution will have to represent Digital Customer Master Data in a centralized, consistent 360-degree view, because neither existing DataWarehouse MIS nor the currently trending Data Lake Architecture can accomplish this task. Master Data Management Administration Architecture will need to be augmented with AI technologies to uncover Digital Customer Master Data. This paper proposes the use of NoSQL MDM Architecture employing an Enterprise Façade API layer for Data Virtualization, Data Standardization, Data Harmonization, and Data Presentation/UI generation. The design decisions taken at various levels of the architecture and implementation of applications will be discussed. The paper will conclude with a brief overview of discussions on researching and developing further AI techniques to implement a production-ready enterprise-level architecture.

References

[1] Pandiri, L. (2023). Specialty Insurance Analytics: AI Techniques for Niche Market Predictions. Available at SSRN.

[2] Koppolu, Hara Krishna Reddy, Goutham Kumar Sheelam, and Venkata Bhardwaj Komaragiri. "Autonomous Telecommunication Networks: The Convergence of Agentic AI and AI-Optimized Hardware." International Journal of Science and Research (IJSR), vol. 12, no. 12, 2023, pp. 2253-2270, https://www.ijsr.net/getabstract.php?paperid=MS2312142403, DOI: https://www.doi.org/10.21275/MS2312142403

[3] Singireddy, J. (2023). Finance 4.0: Predictive Analytics for Financial Risk Management Using AI. European Journal of Analytics and Artificial Intelligence (EJAAI) p-ISSN 3050-9556 en e-ISSN 3050-9564, 1(1).

[4] Lakarasu, P. (2023). Designing Cloud-Native AI Infrastructure: A Framework for High-Performance, Fault-Tolerant, and Compliant Machine Learning Pipelines. Fault-Tolerant, and Compliant Machine Learning Pipelines (December 11, 2023).[

[5] Somu, B. (2023). Towards Self-Healing Bank IT Systems: The Emergence of Agentic AI in Infrastructure Monitoring and Management. American Advanced Journal for Emerging Disciplinaries (AAJED) ISSN: 3067-4190, 1(1).

[6] Somu, B., & Sriram, H. K. (2023). Next-Gen Banking Infrastructure: Designing AI-Native IT Architectures for Financial Institutions. Journal for ReAttach Therapy and Developmental Diversities. https://doi.org/10.53555/jrtdd.v6i10s(2).3610

[7] Gadi, A. L. (2023). Engine Heartbeats and Predictive Diagnostics: Leveraging AI, ML, and IoT-Enabled Data Pipelines for Real-Time Engine Performance Optimization. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 210-240.

[8] Nandan, B. P., & Chitta, S. S. (2023). Machine Learning Driven Metrology and Defect Detection in Extreme Ultraviolet (EUV) Lithography: A Paradigm Shift in Semiconductor Manufacturing. Educational Administration: Theory and Practice, 29 (4), 4555–4568.

[9] Yellanki, S. K. (2023). Bridging the Gap: Aligning Operational Goals with Consumer Behavior via AI-Driven Services. American Journal of Analytics and Artificial Intelligence (ajaai) with ISSN 3067-283X, 1(1).[

[10] Meda, R. (2023). Data Engineering Architectures for Scalable AI in Paint Manufacturing Operations. European Data Science Journal (EDSJ) p-ISSN 3050-9572 en e-ISSN 3050-9580, 1(1).

[11] Burugulla, J. K. R., & Inala, R. (2022). The Future of Payments: A Comprehensive Review of AI, ML, and Cloud Technologies in Finance. Kurdish Studies. https://doi.org/10.53555/ks.v10i2.3870

[12] Sheelam, G. K. (2023). Adaptive AI Workflows for Edge-to-Cloud Processing in Decentralized Mobile Infrastructure. Journal for Reattach Therapy and Development Diversities. https://doi. org/10.53555/jrtdd. v6i10s (2). 3570ugh Predictive Intelligence.

[13] Kannan, S., Annapareddy, V. N., Gadi, A. L., Kommaragiri, V. B., & Koppolu, H. K. R. (2023). AI-Driven Optimization of Renewable Energy Systems: Enhancing Grid Efficiency and Smart Mobility Through 5G and 6G Network Integration. Available at SSRN 5205158.

[14] Annapareddy, V. N., Preethish Nanan, B., Kommaragiri, V. B., Gadi, A. L., & Kalisetty, S. (2022). Emerging Technologies in Smart Computing, Sustainable Energy, and Next-Generation Mobility: Enhancing Digital Infrastructure, Secure Networks, and Intelligent Manufacturing. Venkata Bhardwaj and Gadi, Anil Lokesh and Kalisetty, Srinivas, Emerging Technologies in Smart Computing, Sustainable Energy, and Next-Generation Mobility: Enhancing Digital Infrastructure, Secure Networks, and Intelligent Manufacturing (December 15, 2022).

[15] Raviteja Meda, & Avinash Pamisetty. (2023). Intelligent Infrastructure for Real-Time Inventory and Logistics in Retail Supply Chains. Educational Administration: Theory and Practice, 29(4), 5215–5233. https://doi.org/10.53555/kuey.v29i4.10068

[16] Lahari Pandiri, & Sneha Singireddy. (2023). AI and ML Applications in Dynamic Pricing for Auto and Property Insurance Markets. Journal for ReAttach Therapy and Developmental Diversities, 6(10s(2), 2206–2223. https://doi.org/10.53555/jrtdd.v6i10s(2).3611

[17] Koppolu, H. K. R. Deep Learning and Agentic AI for Automated Payment Fraud Detection: Enhancing Merchant Services Through Predictive Intelligence.

[18] Kalisetty, S., & Singireddy, J. (2023). Agentic AI in Retail: A Paradigm Shift in Autonomous Customer Interaction and Supply Chain Automation. American Advanced Journal for Emerging Disciplinaries (AAJED) ISSN: 3067-4190, 1(1).

[19] Lakkarasu, P. (2023). Generative AI in Financial Intelligence: Unraveling its Potential in Risk Assessment and Compliance. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 241-273.

[20] Somu, B. (2023). Scalable Infrastructure for AI in Banking: Bridging Cloud Computing and Regulatory Demands. Educational Administration: Theory and Practice. https://doi.org/10.53555/kuey.v29i4.10143

[21] Somu, B. (2023). Towards Self-Healing Bank IT Systems: The Emergence of Agentic AI in Infrastructure Monitoring and Management. American Advanced Journal for Emerging Disciplinaries (AAJED) ISSN: 3067-4190, 1(1).

[22] Motamary, S. (2023). Integrating Intelligent BSS Solutions with Edge AI for Real-Time Retail Insights and Analytics. European Advanced Journal for Science & Engineering (EAJSE)-p-ISSN 3050-9696 en e-ISSN 3050-970X, 1(1).

[23] Yellanki, S. K. (2023). Bridging the Gap: Aligning Operational Goals with Consumer Behavior via AI-Driven Services. American Journal of Analytics and Artificial Intelligence (ajaai) with ISSN 3067-283X, 1(1).

[24] Motamary, S. (2023). Enhancing Retail Infrastructure Agility through Intelligent OSS and ML-Orchestrated Workflows. Available at SSRN 5272164.