

zensar

Leveraging AI to Clean, Organize, and Optimize Customer Registries for Increased Sales and Marketing



An  **RPG** Company

Executive summary

Master Data Management (MDM) is a comprehensive method to define and manage an organization's critical data for a unified, single source of truth across all functions. The process offers a 360-degree view of data and enables users to deliver better business insights through self-service analytics. The need for MDM primarily arose from the need for businesses to enhance productivity and customer experience by improving the quality and consistency of data, including product, asset, and customer data.

According to one research, the MDM market size is expected to grow from \$16.7 billion in 2022 to \$34.5 billion by 2027. This growth will be driven by key industry challenges and trends that will influence and speed up the adoption of MDM. These include:

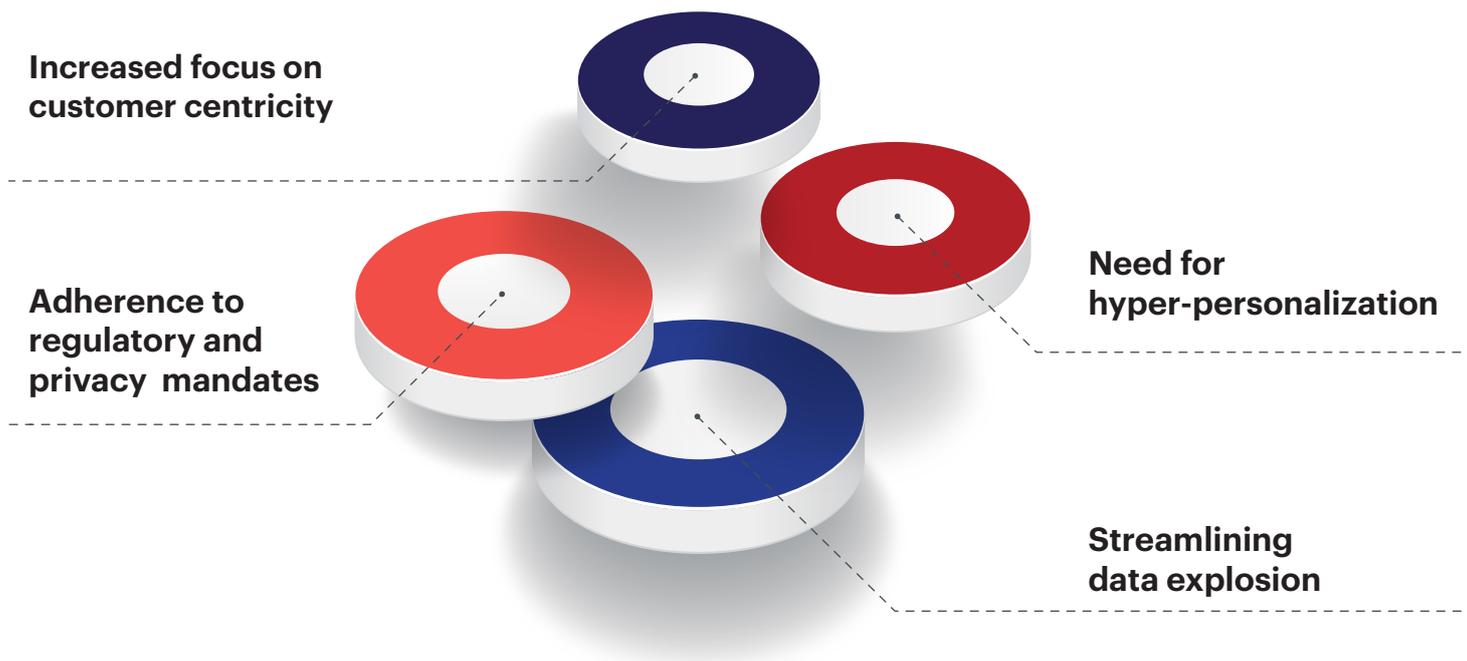


Figure 1: Factors influencing the adoption of MDM

Why MDM

Last year I switched jobs after 16 years and moved my corporate mobile number to my personal account, hoping that I wouldn't have to share my contact information again with all the hundreds of colleagues, family, and friends. But each time I called a new person, they saw my caller ID as my previous organization. I contacted the carrier to fix this and even changed my carrier. But the issue remained. Frustrating, right? Such data issues are a common occurrence across industries. And often,

only because applications across large enterprises don't subscribe to the master data. The result is poor customer experience, revenue loss, and impact on service.

Achieving data accuracy, consistency, and control is critical as organizations become more dependent on data for all their day-to-day operations. When executed effectively, MDM can help organizations:

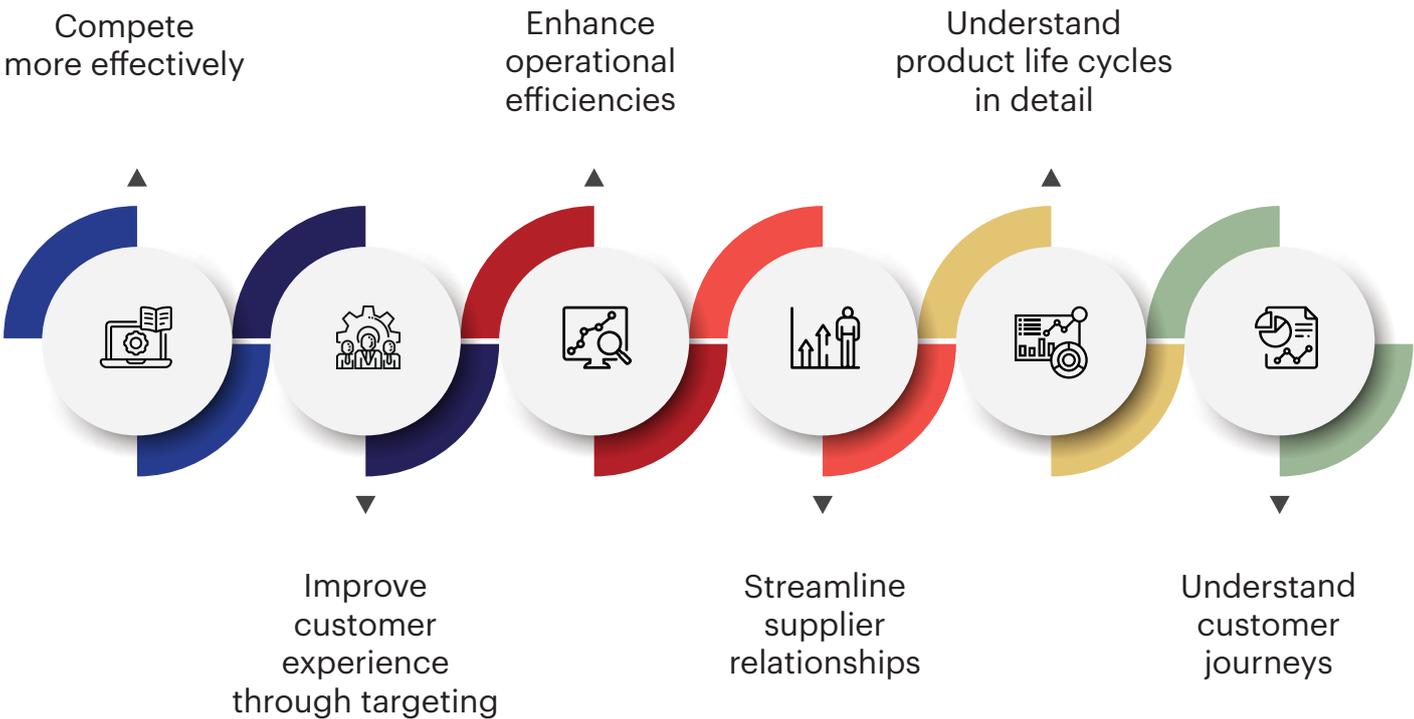


Figure 2: Advantages of an MDM

Here's a quick look at how well-architected MDM can benefit different verticals:



Healthcare providers can access patient data faster for diagnostics and treatments



Banking and financial organizations can reduce customer churn by providing timely, accurate, and personalized services



Insurance companies can improve claims processing efficiency



Energy companies can balance supply and demand more accurately



Supply chains can reduce waste (warehouse space, fuel, etc.)



Retailers can synchronize online and brick-and-mortar channels



A simple scenario of why enterprises need effective MDM

A customer downloads a white paper, and their contact information is stored in the marketing database. The customer later requests a quote, and this information is captured in the sales database. Finally, the customer places an order, and that data is stored in the finance team's database. In

such a scenario, there is an urgent need to create a master record of customers for the enterprise to be aware that it is working with the same individual. This is a big step toward improving customer experience and providing a seamless service.



Figure 3: Master data as a single source of truth

Understanding types of MDM systems

Typically, two forms of MDM are implemented separately or in tandem. These are **analytical MDM**, which aims to feed consistent master data to data warehouses and other analytics systems, and **operational MDM**, which focuses on the master data in core business systems. Both provide a systematic approach to managing master data, typically enabled by deploying a centralized MDM hub where the master data is stored and maintained.

However, there are different ways to architect MDM systems, depending on how organizations want to structure their master data management programs and the connections between the MDM hub and source systems. The primary MDM architectural styles that have been identified by data management consultants and MDM software vendors include the following:



Registry architecture

A registry architecture that creates a unified index of master data for analytical uses without changing any data in individual source systems. Regarded as the most lightweight MDM architecture, this style

uses data cleansing and matching tools to identify duplicate data entries in different systems and cross-reference them in the registry.





Consolidated architecture

A consolidation approach, in which sets of master data are pulled from various source systems and consolidated in the MDM hub. This creates a centralized repository of consistent master data primarily used in BI,

analytics, and enterprise reporting. But operational systems continue to use their own master data for transaction processing.



Co-existing architecture

A coexistence style that creates a consolidated set of master data in the MDM hub. In this case, changes to the master data in individual source systems are updated in the hub and can then be

propagated to other systems, so they all use the same data. This offers a balance between system-level management and centralized governance of master data.



Centralized architecture

A transaction or centralized architecture approach moves all management and updating of master data to the MDM hub, which publishes data changes to each source system. It is the most intrusive style

of MDM from an organizational standpoint because of the shift to complete centralization. However, it is also the one that provides the highest level of enterprise control.

Master data management grew out of previously separate methodologies focused on consolidating data for specific entities — customer data integration (CDI) and product information management (PIM). MDM brought them together into a single category with a broader focus, although CDI and PIM are still active subcategories.



How is MDM evolving

Various factors, such as the increased use of data quality tools for data management and the rising need for compliance, are creating the need for MDM solutions to evolve.



Data governance platforms

Data governance platforms are becoming more sophisticated, providing more advanced capabilities such as data lineage tracking, data discovery, and data impact

analysis. This allows organizations to better understand how data flows through their systems and identify and resolve data issues more quickly and effectively.



Data quality tools

The increased data churn necessitates sophisticated data quality technologies to improve operational excellence. Data quality is the first step toward implementing MDM, a powerful tool that allows businesses to trace the stages of information flow and

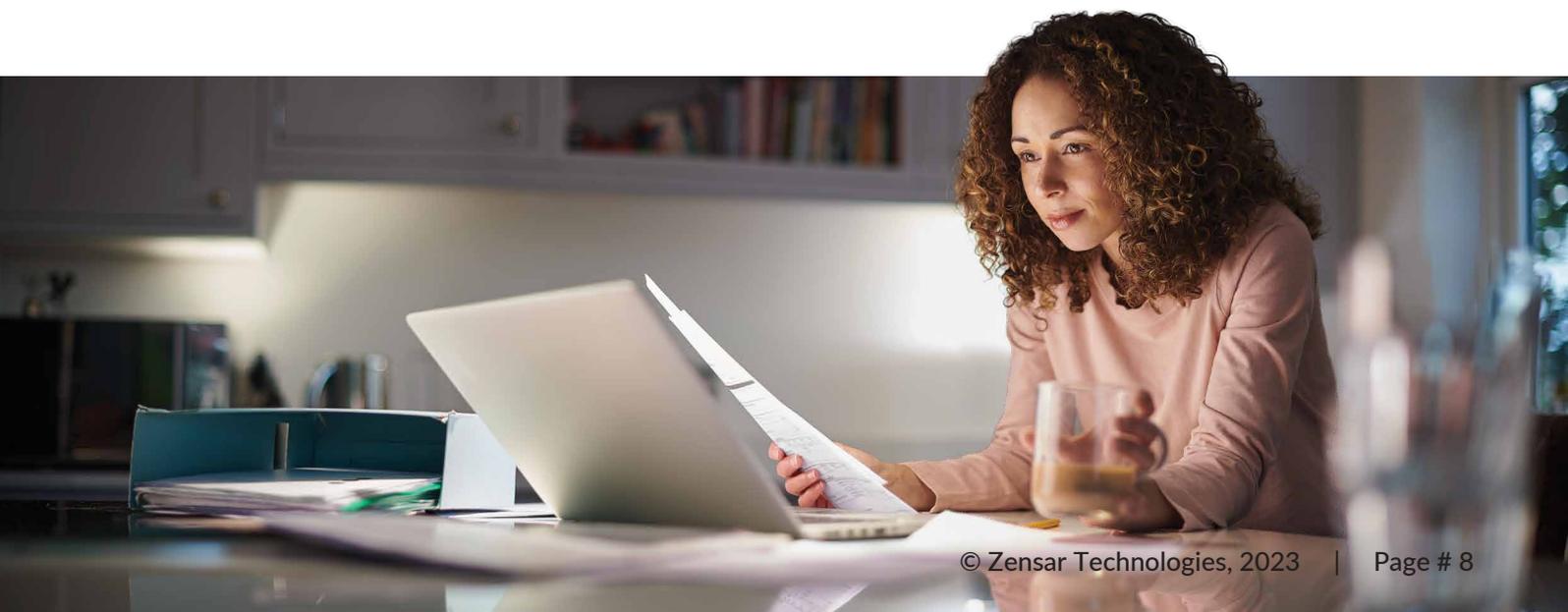
document data transformation. Data quality technologies assure data accuracy, while MDM solutions automatically update records and clear superfluous data. This has allowed businesses to combine the adoption of MDM solutions and data quality technologies.



Data integration

Data integration tools are becoming more flexible and can now integrate data from a broader range of sources, such as social media, IoT devices, and cloud-based

services. This allows organizations to gain a complete view of their data and use it for improved insights and decision-making.





Artificial intelligence (AI), big data, and machine learning (ML)

These new-age technologies provide advanced technical capabilities for data storage and processing, as well as ensuring huge dataset availability. Enterprises have a broad range of requirements as new tech capabilities in managing multi-domain and multi-faceted data have improved. Common master data challenges, such as duplicate records, are being identified, and data quality is assured. The use of ML in MDM is growing, allowing fresh data to be

considered while relying on previous data exchanges. Organizations can use ML to detect data patterns and propose linkages, correlations, and speedier data adaption from a source to a consumer. Enterprises can integrate the source for CRM, personalization, social commerce, and offline activity with the help of these technologies, resulting in a single data source for all channels.



Cloud-based solutions

Cloud-based MDM solutions are becoming more popular as they provide a centralized, scalable, and cost-effective way to manage and maintain data across an organization. They can be accessed from anywhere and

easily integrated with other cloud-based tools and services, such as data integration, quality tools, and data governance platforms.





Data catalog

More sophisticated data catalogs can now provide advanced capabilities such as semantic search, data lineage tracking, and data impact analysis, enabling organizations

to understand their data flow and identify and resolve data issues more quickly and effectively.



Blockchain

Blockchain has become mainstream and is being used across industries to ensure data integrity and immutability. This is essential to facilitate secure data sharing and collaboration across organizations. The distributed nature of blockchain technology

allows multiple parties to have access to the same data, and once recorded to the blockchain, data cannot be altered. This makes it particularly useful for MDM in industries where data integrity is critical, such as finance and healthcare.

To avoid damaging their brand, prominent enterprises combine MDM with other modern technologies to provide complete, consistent, and accurate information across all channels.



How is Zensar helping clients manage data effectively with MDM and AI

Many of our enterprise clients deal with a large number of customers. Some of these enterprises have millions of customers whose data must be managed to provide regular services. However, data management involves much complexity and requires significant time and money.

Challenges in managing data



Data is often arranged hierarchically to ensure that the subsidiaries are treated the same way as the parent organization. These enterprises keep merging/demerging and are also sold/bought, making the data noisy often.

Organizations generally employ a large number of people to clean the data manually. Further, these teams do not keep the traceability of the source, which made them change the information in the original database.

Finally, scaling up team size becomes problematic when this data has to be cleaned in large volumes due to external demand.

To overcome all these problems, Zensar has developed an AI-based offering called Customer Registry Cleanup Framework (CRCF) to continuously denoise data.



Customer Registry Cleanup Framework (CRCF)

CRCF uses open-source documents and many regulatory documents to extract the information required to update the database. We use a knowledge graph-based approach to maintain data integrity. This data can be collected using various automation frameworks from the web.

Knowledge management technologies have evolved to an extent where large databases can now be represented graphically to perform suitable queries and updates and store them with the ability to update graphs when new information is discovered.

Knowledge graphs need the support of many involved natural language processing (NLP) technologies such as Named Entity Recognition, Named Entity Disambiguation, Iterative Merging, and more. Various entities like organizations, key people, industries, regions, and any related entities can be discovered and mapped to one of the nodes in the knowledge graph. Since interpreting text into a map of one of the many entities listed above requires entity recognition, the transformer-based method uses state-of-the-art NLP techniques to identify entities.



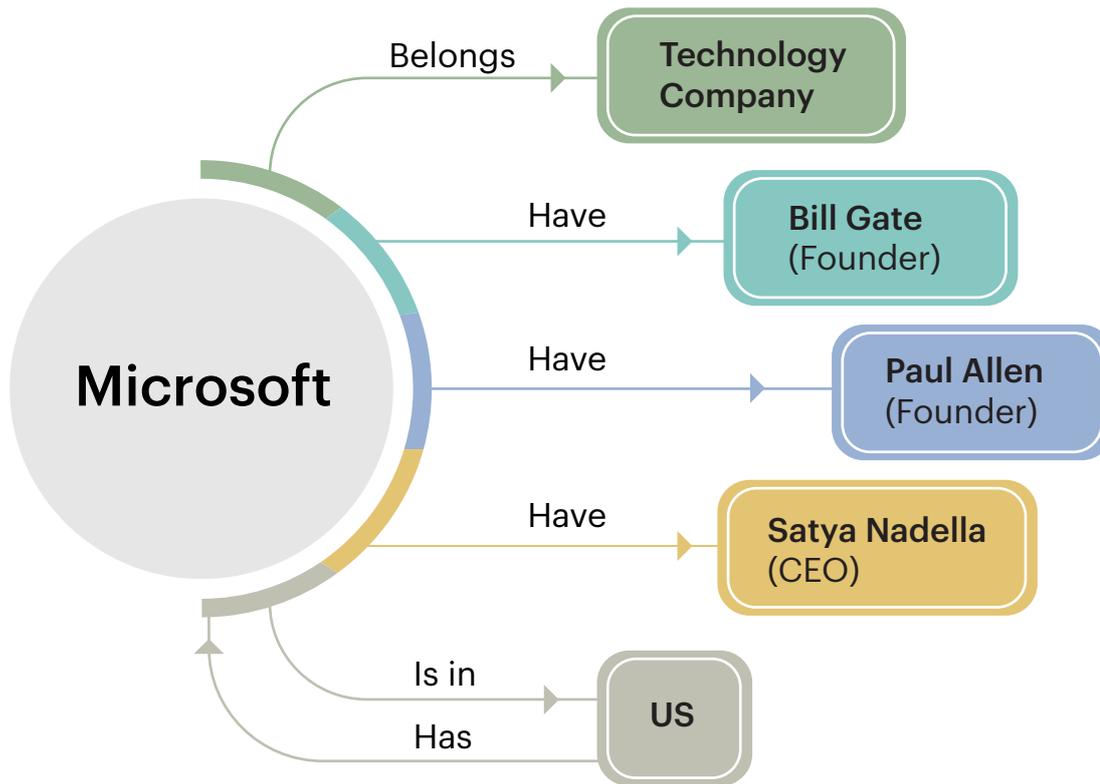


Figure 4: A sample of an organization and its relevant information in the knowledge graph

Once the entities are identified, it is easier to understand the relationship as represented in a sentence. Simple techniques like triplet extraction can be utilized using dependency parsing techniques. More involved (deep learning-based) methods using BERT can also be used for the same task.

Zensar uses a second stream of advanced NLP techniques (deep learning-based) to keep the sanity of the knowledge graph for the disambiguation and linking of identical entities. Many entities can be represented in alternate forms, which need to be detected, logged, and merged to keep common knowledge of each entity. There are multiple steps to be followed, such as

candidate entity selection, contextual entity shortlisting, and disambiguation of the entity. With all the advances in NLP, we cannot do away with humans in the loop to verify low-confidence entity disambiguation and linking.

Appropriate business metrics must be captured to track the success of knowledge management models judiciously. Beyond the model's accuracy, detecting that relationships between entities have changed is of primary importance. Hence, minimizing the false negatives can be seen as taking precedence over false positives.

Advantages of CRCF

Using CRCF offers several critical benefits to the organization, including:



Continuous cleanup of customer registry



Improved sales and marketing activities



Traceability of decision for each cleanup



High-quality outcomes due to a larger volume of data processed per request



Fast scale-up and scale-down to accommodate custom requests



Faster downstream tasks that deliver higher value to the organization



Faster downstream tasks that deliver higher value to the organization



100 percent traceability as we can keep track of the data update



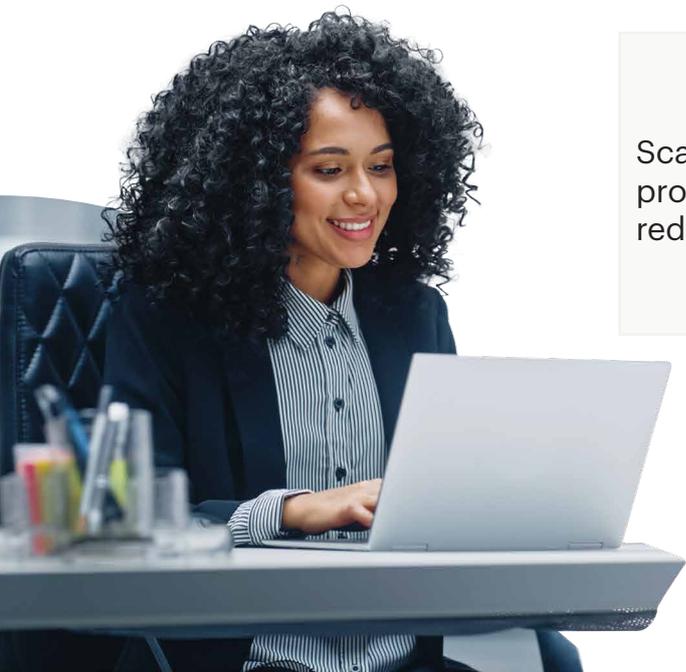
Non-linear increase in capacity of manual cleanup operations



Scalable and flexible processes with reduced cost



Enhanced data quality with the overall reduction in registry-related challenges



Conclusion

There is no doubt that MDM is a robust method for defining and managing an organization's critical data for a unified, single source of truth across all functions. A 360-degree view of data enables users to deliver better business insights through self-service analytics and enhance productivity and customer experience by improving the quality and consistency of data, including product, asset, and customer data. The MDM market will grow in the coming years, given the increased focus on customer centricity, the need for hyper-personalization, adherence to regulatory and privacy mandates, and streamlining data explosion.

With CRCF, Zensar has helped clients achieve benefits such as increased process efficiency, reduced time and cost of operations, along with the other benefits mentioned in the paper. A well-architected MDM that utilizes different technologies to create contextualized solutions, such as CRCF, can help organizations compete more effectively. This is possible as MDM helps improve customer experiences, enhance operational efficiencies, streamline supplier relationships, and understand customer journeys and product life cycles in detail.

Authors

Abhijeet Vaidya

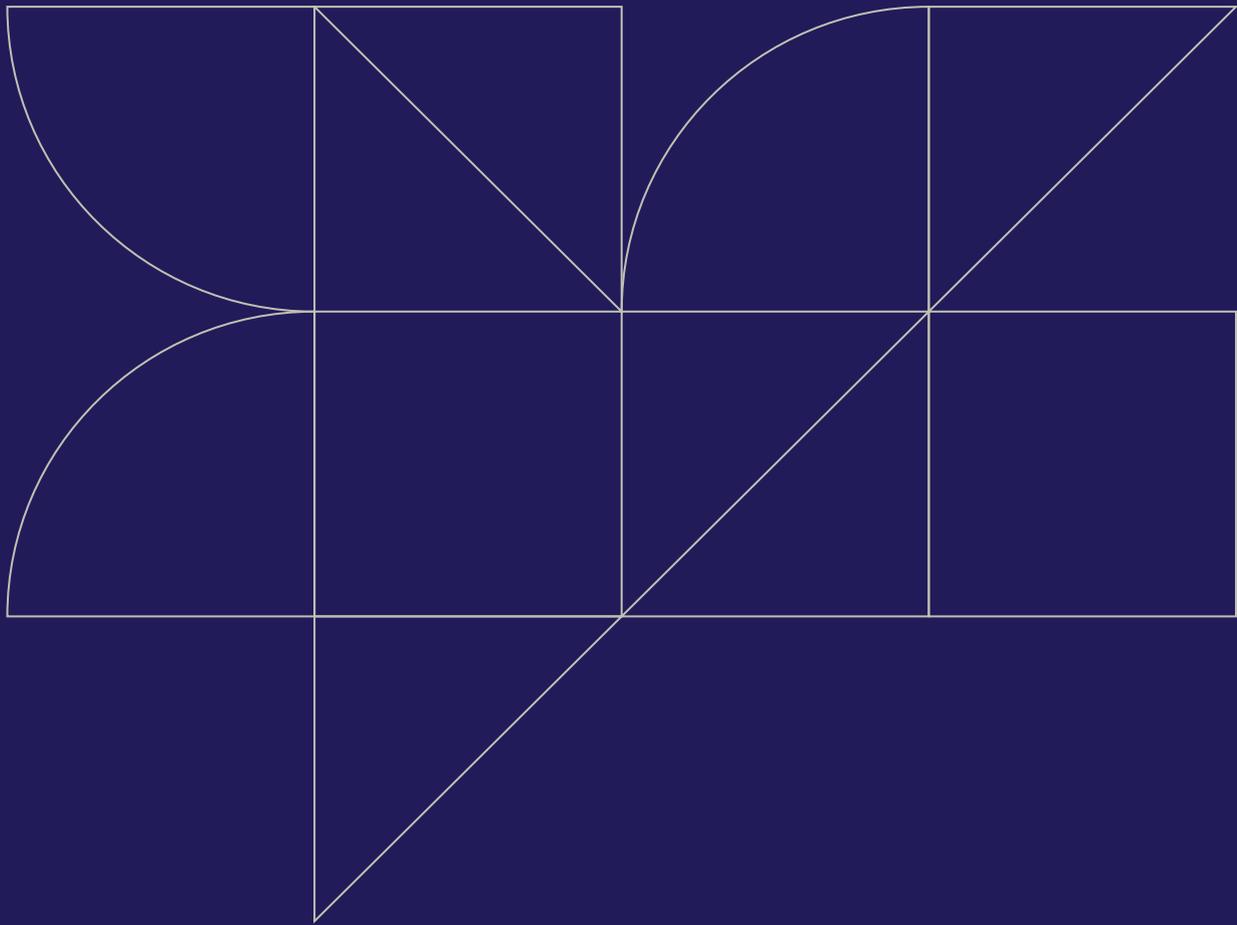
Head - Global Markets, Data Engineering and Analytics
abhijeet.vaidya@zensar.com

Dr. Sumant Kulkarni

AVP and Head - Artificial Intelligence and Machine Learning
sumant.kulkarni@zensar.com

Dr. Kumar Dheenadayalan

Manager - Data Science
k.dheenadayalan@zensar.com



zensar

An  RPG Company

We conceptualize, build, and manage digital products through experience design, data engineering, and advanced analytics for over 145 leading companies. Our solutions leverage industry-leading platforms to help our clients be competitive, agile, and disruptive while moving with velocity through change and opportunity.

With headquarters in Pune, India, our 11,500+ associates work across 30+ locations, including Milpitas, Seattle, Princeton, Cape Town, London, Singapore, and Mexico City.

For more information please contact: velocity@zensar.com | www.zensar.com