

Data Modernization

Enable Real-time Data-driven Decisions
at a **Faster Speed**

The recent disruptions triggered by breakthroughs in digital transformation and technological advancements in cloud computing, data analytics and artificial intelligence led to a massive surge in data. It eventually snowballed and became a double-edged sword as businesses were grappling with the data deluge from various sources and the escalating cost of data storage. Can businesses move forward with siloed information in the age of AI, accelerating cloud adoption and digital disruption? The days of siloed data being narrowly defined as highly structured information from a few specific sources are long gone. Consider the consequences if businesses fail to keep pace with the increasing operational and analytical data requirements, impeding their ability to make quick decisions and gain critical insights in real-time. Today, as businesses are embracing the data driven culture to remain ahead of the curve, it is critical to identify the stakeholders for data usage, ascertain the purpose behind using the data and establish a process of democratization of data across the entity.

“Data morphosis” through data modernization is the reality of organizations looking to unify the various data types from different sources. Will the transformation journey be smooth ? Like any other new

adoption, we are bound to confront barriers and resistance in delivering data forward solutions. In light of the myriad uncertainties, it is equally important to evaluate the best practices in building a Modern Data Platform while treading with a cautious approach in the modernization journey. Data modernization is driving the adoption of new age data platforms, referred to as Modern Data Platforms, path breaking architectures and approaches backed with robust data protection features ensuring sensitive data security. This paper by Everest Group, aims to provide an insight on the significance of data modernization as a continuous process, to provide a unified and scalable platform, enable real time data driven decisions at a faster speed and attain a competitive edge in the rapidly evolving digital landscape.

Rajsekhar Datta Roy

Chief Technology Officer





Data Modernization: Enable Real-time Data-driven Decisions at a Faster Speed



This document has been licensed to **Sonata Software**

Vishal Gupta, Vice President
Priya Bhalla, Practice Director

Copyright © 2023, Everest Global, Inc. All rights reserved.

www.everestgrp.com

Contents

Introduction	03
The importance of an MDP	04
What is an MDP?	05
Emerging paradigms for building an MDP	07
Enterprise adoption of MDPs	12
Best Practices for building an MDP	15
Case study	17
Future outlook	18

Introduction

Organizations today face an unprecedented deluge of data, which holds immense potential to drive innovation, uncover valuable insights, and facilitate informed decision-making. To fully leverage the potential of data, enterprises are adopting Modern Data Platforms (MDPs) that provide increased scalability, flexibility, and security. The exponential growth of data sources, formats, and analytical tools renders traditional data management and governance approaches obsolete. As a result, organizations grapple with critical challenges such as data silos, escalating data management costs, and suboptimal data quality, resulting in limited insights.

An MDP not only lays the foundation for data-driven decision-making but also helps organizations overcome the challenges posed by traditional data platforms. It serves as a central nervous system, empowering organizations to efficiently ingest, store, process, and analyze vast and diverse datasets. In this report, we will explore the transformative potential of MDPs in data management practices, delving into the broader scope of MDPs and their impact on organizations' ability to harness the true value of data. Additionally, we will examine three innovative data management practices – data fabric, data mesh, and data vault – that enable organizations to address the complexities of data and achieve true data democratization.

Data fabric represents a federated approach to data integration, serving as a unified layer that seamlessly links diverse data sources, applications, and analytics tools across the entire organization. Conversely, data mesh adopts a decentralized approach to data management, recognizing data as a product to be owned and managed by individual domain teams. Furthermore, we will explore the merits of data vault within the context of MDPs. Data vault is a data modeling architecture that provides a scalable and agile foundation for data warehousing. It emphasizes data traceability and scalability, achieved by segregating core business data from historical data.

This viewpoint serves as a valuable resource for enterprises seeking to navigate the complexities of traditional data platforms by adopting an MDP. It provides insights on how to build scalable, self-serve data ecosystems, transforming data into a strategic asset rather than a liability. The report will educate enterprises about emerging data management paradigms and essential best practices to consider before embarking on their MDP journeys.

The importance of an MDP

In 2022, enterprises faced numerous uncertainties and disruptions, with notable technological advances like ChatGPT disrupting the market while also contending with global supply chain gaps, geopolitical crisis, and financial instabilities. Amid these challenges, one constant for enterprises was the surge in digitization across all fronts. According to the World Bank¹, the digital economy contributes 15% to global GDP and has been growing 2.5 times faster than the physical world GDP in the past decade.

The widespread adoption of digital technologies has facilitated the collection and storage of vast amounts of data. Advances in cloud computing, data analytics, and Artificial Intelligence (AI) have made it easier to process and analyze large data volumes in real time. As a result, every process, workflow, or interaction within a modern enterprise now leaves a digital trace or has a complete digital counterpart. This has led to a significant increase in data generation and accumulation by enterprises.

In 2023, the data-driven trend is expected to continue and even intensify. We generate over two quintillion data every day, with no signs of slowing down. Compounding the complexity, data now takes diverse forms, ranging from traditional text, audio, video, and images to real-time data, multidimensional data, and synthetic data. The primary responsibility of any data leader is to ensure the availability of the right form/type of data to the appropriate individuals, whether they are business teams, data scientists, or analysts, precisely when they need it the most. In other words, the focus is on solving the following aspects:

- Identifying who is using the data
- Understanding what they are using it for
- Establishing ways to democratize data across organizations

As simple as it may sound, fulfilling these objectives is an arduous task that demands enterprises to rethink and modernize their data strategies. Enterprise leaders find themselves grappling with the data dilemma: while they aspire to achieve ambitious digital transformation goals, their traditional, monolithic data infrastructure presents itself as their biggest obstacle. In essence, the three major challenges data leaders currently face are:

- **Siloed data and extensive data replication, resulting in soaring costs:** In many leading enterprises, data is scattered across fragmented data warehouses spanning multicloud and on-premises infrastructure. Operational and analytical systems often operate independently and feed into different data platforms. Consequently, when a business leader or data scientist requires access to specific data, they must resort to outdated methods of collating data from various warehouses, leading to repetitive data replication. This process incurs high infrastructure and storage costs while increasing overall complexity. Data leaders are increasingly recognizing that the current approach fails to meet the growing demands for operational and analytical data, necessitating a unified view of customers and real-time insights
- **Centralized control and management of data, which hinders true data democratization:** As data has proliferated across enterprises, multiple disconnected data platforms have emerged, each managed by separate data management teams lacking the necessary business context. Business users increasingly require real-time access to data to extract valuable insights without heavy reliance on IT or central data teams. The centralized management of data poses a considerable obstacle

1. [World Economic Forum](#)

for data leaders, emphasizing the need for self-serve, automated data platforms that empower business users to access and utilize data independently

- **Poor data quality, resulting in erroneous business insights:** Traditional data warehouses and data lakes are often connected to multiple data sources without consistent quality standards or centralized data governance. This lack of common standards hampers the utilization and sharing of data among different units within the organization, impeding seamless data interoperability and hindering the creation of reliable insights

As enterprises face mounting complexities, there is a pressing need to modernize traditional data platforms and replace legacy data architectures with next-generation structures that promote data-enabled decision-making.

What is an MDP?

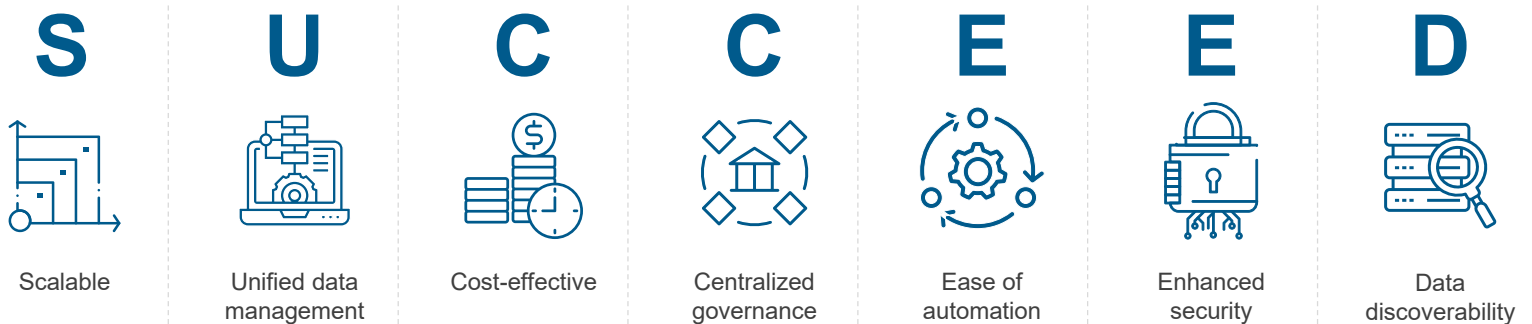
Over time, data platforms have evolved to adapt to the evolving needs of enterprises and the data landscape. Initially focused on collecting customer and machine-generated data, they have now transformed into comprehensive solutions that encompass the entire data life cycle. This continuous innovation has given rise to MDPs, which are highly advanced, agile, and efficient compared to their traditional counterparts. **An MDP is a unique assortment of technologies, techniques, and services, stitched together into a flexible, secure, and scalable platform, to help enterprises advance their data analytics and AI initiatives.**

Troubled by the daunting challenges of traditional data platforms, data leaders are swiftly adapting their strategies to build a truly democratic data platform. This platform will fulfill the real-time data requirements of users through automated self-service capabilities. Exhibit 1 illustrates next-generation data platforms' distinct characteristics that set them apart from their traditional counterparts and enable enterprises to **SUCCEED** in their digital transformation journeys.

EXHIBIT 1

Characteristics of an MDP

Source: Everest Group (2023)



Below, we discuss these characteristics in detail.

- **Provide scalability** – to manage the data deluge: Unlike traditional data platforms where the same compute capacity handles both data ingestion and query functions, MDPs decouple storage and compute functions. This decoupling enables the data stack to scale automatically, providing elasticity and accommodating enterprises' complex data demands, thereby mitigating performance and latency issues associated with real-time analytics
- **Unified data management** – to tackle the challenge of data siloes: Traditional data storage mechanisms such as databases, data warehouses, and data lakes allowed enterprises to centralize their data. However, as data volumes skyrocketed, isolation led to obsolete and fragmented data silos. A key feature of MDPs is their ability to unify data sources and facilitate seamless data movement across systems. This is typically achieved through data virtualization techniques, where data integration occurs without physical replication. By minimizing delays and latencies, MDPs enable efficient data access and integration across the organization, mitigating the issues associated with data silos
- **Offer significant cost savings** – to address the rising IT costs: Unlike traditional platforms that require expensive investments in building and managing on-premises infrastructure, most MDPs are cloud-based and offer easy setup options. This eliminates the need for enterprises to spend millions on infrastructure. Additionally, cloud-based data platforms operate on a consumption-based, pay-as-you-go pricing model, which allows enterprises to optimize their data costs by paying only for the resources they use, resulting in substantial cost reductions
- **Centralize governance** – to improve data quality: While it may result in longer cycle times, some level of centralization is necessary to establish data standards and policies, reducing the risk of errors and inconsistencies in data analysis. A centralized governance approach ensures a structured management of data assets across an organization, promoting consistency and accuracy. In the next phase of digital transformation, where data plays a vital role, having controls in place for data content, structure, use, and safety becomes paramount. Enterprises aspire to have a seamless flow of reliable, real-time data across the organization and even with third parties. Traditional top-down approaches to data governance are often ineffective and non-scalable in such scenarios. An MDP reimagines data governance as a democratic responsibility rather than a standalone function. It embraces the principles of centrality and responsibility, providing a centralized, scalable, and automated approach to data governance. This ensures regulatory compliance and effectively manages data at every stage of its life cycle
- **Ease of automation** – to boost self-service: MDPs prioritize automation to facilitate self-service capabilities, a key factor in achieving true data democratization. As organizations grow in size and data volume, relying on a centralized data team for every small task becomes impractical and burdensome. To enable real-time data access for every member of the organization, MDPs leverage automation and self-service features. These empower data users to explore, visualize, and search for information without the need for highly skilled data engineers
- **Enhanced security** – to protect sensitive data: MDPs prioritize enhanced security to safeguard sensitive data and comply with data protection regulations. These platforms incorporate data-level security measures, ensuring the protection of confidential information. AI-driven access management features enable intelligent access controls based on user roles, permissions, and privileges, limiting access to sensitive data. Furthermore, advanced data observability tools provide real-time continuous

monitoring, enabling organizations to promptly detect and respond to potential threats or vulnerabilities. By prioritizing security, MDPs establish a safe and secure data environment, instilling confidence in data users and ensuring compliance with privacy regulations

- **Easy data discoverability** – to achieve true data democratization: MDPs do not treat data as a by-product of digitization but as a valuable product that accelerates digitization efforts. By organizing data as standardized products based on consumption archetype patterns, MDPs facilitate efficient data access and reduce the time required for data consumers to derive value. This approach ensures that data is readily available to drive real-time insights and support informed decision-making throughout the organization and beyond. With easy data discoverability, MDPs empower enterprises to unlock the full potential of their data assets and maximize their impact on business outcomes

In light of these attributes, enterprises are reevaluating their data strategies and building MDPs. An MDP serves as a comprehensive architecture designed to securely and efficiently collect, process, store, and analyze large volumes of data from diverse sources. It encompasses various components, including data ingestion tools, data processing engines, data storage solutions, data integration tools, data quality and governance tools, and analytics and reporting tools.

This functional overview highlights that a modern data platform is not simply a singular tool or platform, but rather a complex architecture that seamlessly integrates multiple components to enable effective data management and analytics. In recent years, there have been significant changes in the field of data management driven by the exponential growth of data, the widespread adoption of cloud, and the increasing use of AI. These factors have necessitated the evolution of traditional data management approaches to meet the demands of modern data environments. Data management has shifted toward more decentralized, autonomous, and scalable data architectures to address the evolving needs of enterprises. The following section of this report will delve into emerging data management paradigms such as data fabric, data mesh, and data vault, which have garnered considerable attention from enterprises seeking innovative solutions in the realm of data management.

Emerging paradigms for building an MDP

As mentioned earlier, an MDP is a comprehensive platform that combines various technologies, techniques, and services to create a flexible, secure, and scalable solution. Every organization has its own unique requirements, and in a rapidly advancing technological landscape, there is no one-size-fits-all approach to building an MDP. Exhibit 2 highlights the emerging data management architectures and approaches that play a vital role in developing a MDPs. These innovative solutions have successfully passed the commercial viability stages and are delivering significant benefits to enterprises as they embark on their data modernization journeys.

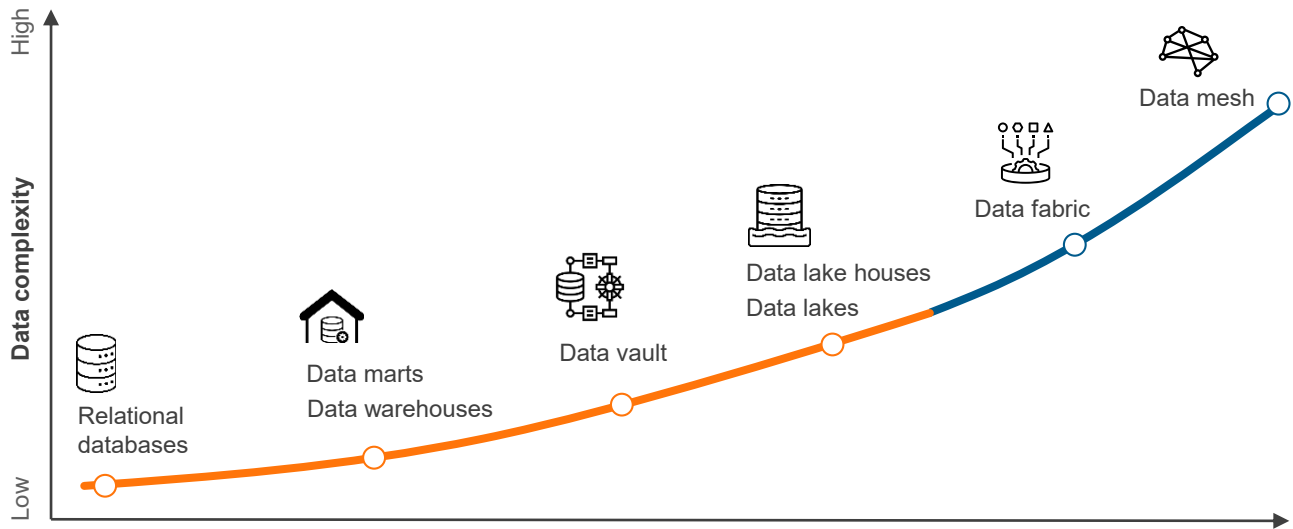
An MDP stitches together a unique assortment of technologies, techniques, and services into a flexible, secure, and scalable platform that helps enterprises advance their data analytics and AI initiatives.

EXHIBIT 2

Evolving data management approaches

Source: Everest Group (2023)

- Era of data consumption – data as a by-product of digitization
- Data as product – a catalyst for digitization



1990s

- Reporting as a driver
- Uses data to solve targeted use cases
- Low/No data security

Early 2000

- Analytics as a driver
- Uses data to solve BU level problems
- Infrastructure-level security standards introduced

2018 and beyond

- Insights as a driver
- Data-driven business decisions
- Regulatory push toward data-level security

Data fabric – standardizing data management across environments

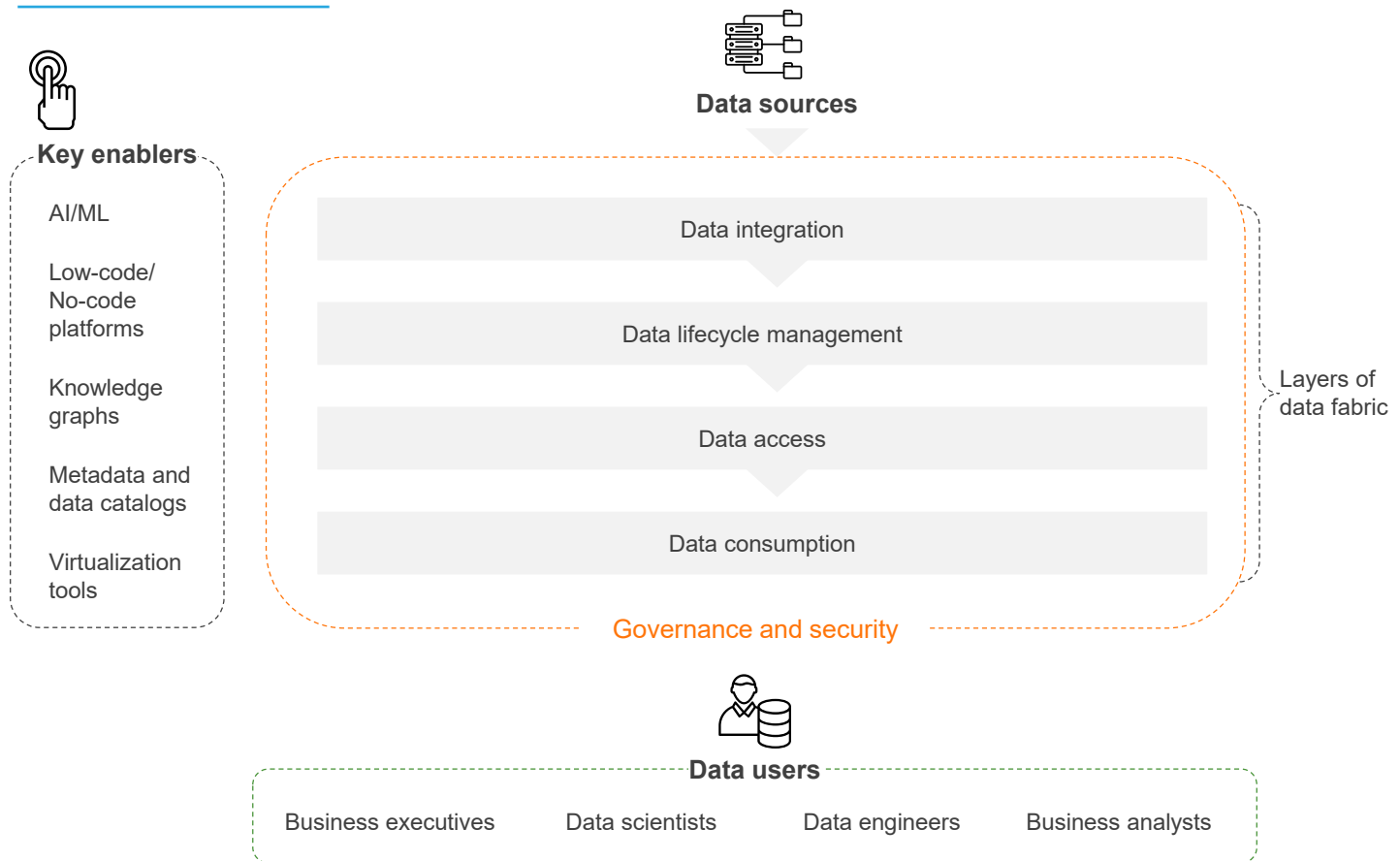
The data fabric architecture, depicted in Exhibit 3, unifies the real-time management, governance, and data access across multiple environments, including cloud, on-premises, and edge devices. It achieves this by leveraging standard tools and processes to provide a cohesive and seamless data experience.

“ We can enable eClinical Solutions’ customers, who are drug companies or companies involved in drug exploration, to access data across separate data stores, distributed data clouds, data lakes, or data warehouses. This is where data fabric can help.

– Raj Indupuri, Chief Executive Officer, eClinical Solutions

EXHIBIT 3**Data fabric architecture**

Source: Everest Group (2023)



The typical data fabric architecture comprises four layers that work together to enable seamless data management and access.

- **Data integration layer:** This layer, often implemented as a virtualization layer, facilitates data integration from various sources without the need for physical replication. By utilizing data services and APIs, the data fabric simplifies the data ingestion process, allowing data to be pulled from diverse systems such as data lakes, data warehouses, and databases. This virtualization layer abstracts the complexities associated with individual systems, providing users with easy access to integrated data
- **Data lifecycle management layer:** This layer handles data transformation, ensuring data usability, and manages data quality. It also monitors data bias, enforces data governance and security principles, and governs the datasets accessed across different environments. By taking care of these aspects, the data fabric ensures that data is reliable, accurate, and compliant with regulations
- **Data access layer:** This layer connects to multiple environments through adapters or connectors, using a low-code/no-code approach. This empowers users to access data in a self-service manner, eliminating the need for technical expertise or constant involvement from IT teams. The layer enables seamless data access across various environments, ensuring that users can retrieve the data they need efficiently

- **Data consumption layer:** This layer includes analytical and AI applications that require access to datasets. It may include features such as data cataloging and searchable interfaces, simplifying the process of discovering and accessing relevant data. The data consumption layer facilitates easy data exploration and analysis, enabling users to derive insights and drive decision-making

Data fabric revolutionizes data access by seamlessly integrating it into applications and business workflows in real time, offering data users a comprehensive view of enterprise-wide data. By unlocking siloed data, data fabric unlocks new business opportunities, empowers data teams, and enables a truly data-driven enterprise. It addresses the limitations of traditional data management approaches, particularly in managing real-time unstructured data at scale. The virtualization layer empowers enterprises to effortlessly add new data sources and applications to their data platforms, ensuring scalability and flexibility.

To work effectively, a data fabric relies on several tools such as data ingestion tools, virtualization tools, cloud computing, data governance and security tools, and AI. While these tools form the foundation of a data fabric, technologies such as knowledge graphs are gaining popularity for managing and integrating data within the data fabric. Knowledge graphs use semantic relationships to enable meaningful data search and querying. By tracing data lineage and provenance, knowledge graphs provide users with valuable insights into the history and context of the data they are working with.

The primary value proposition of data fabric lies in its ability to facilitate seamless data movement. While initial implementations focused on unifying data across multiple clouds, the true potential of data fabric lies in unifying data across hybrid infrastructures, both within and beyond enterprise boundaries. By enabling data consolidation and integration, data fabric empowers organizations to maximize the value of their data assets and accelerate their digital transformation initiatives.

Data mesh – rising popularity of the data product mindset

Data mesh represents a paradigm shift in data management by adopting a decentralized approach. Unlike traditional centralized data management, data mesh entrusts the ownership and governance of data to the domain experts within the organization who possess a deep understanding of the data. It is an emerging concept that enterprises are increasingly embracing to define their operating models in distributed data environments. The approach governs how the organization manages and accesses disparate datasets owned by different business functions. It is based on the concept of decentralized data ownership and involves a strategic shift in the organizational mindset.

Below, we look at how the components of a data mesh are relevant for modern enterprises:

- **Distributed ownership under functional domains:** Data mesh recognizes the value of domain expertise and encourages the ownership of data by respective functional teams. This approach ensures that teams take full responsibility for their data, including its quality, design, documentation, and maintenance. In a multicloud scenario, distinct functions within an enterprise may have varying Data, Analytics, and Artificial Intelligence (DAAI) requirements. Data mesh provides them with the flexibility to choose their own cloud provider based on their specific data needs
- **Data-as-a-product approach:** Data mesh promotes the concept of treating data as a product. Functional teams take ownership of and manage the data, ensuring the development of data products that facilitate the sharing of insights across other functional teams. This promotes enhanced collaboration and prevents the formation of data silos within the organization that may rise from functional ownership of cloud data and the adoption of multiple clouds by different teams

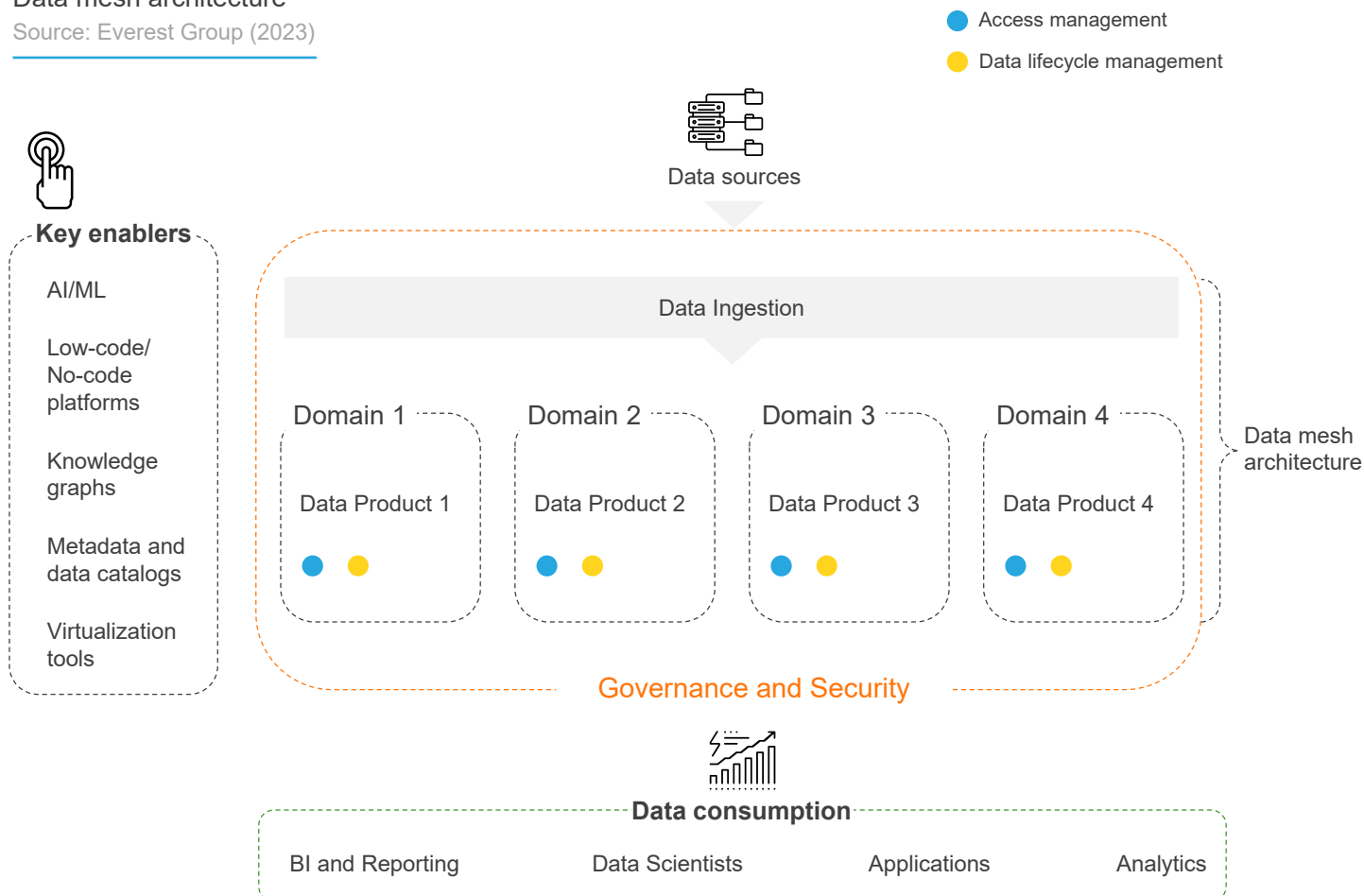
- Self-service architecture:** Enabling self-service access to datasets is a vital requirement for both business users and data professionals, as it allows for easy and direct access to relevant data, facilitating insights generation and provisioning for application development and testing. By implementing a self-service data platform, organizations ensure seamless interoperability between functional datasets and applications across various environments, enabling faster access for business users who no longer need to rely on data professionals to extract usable datasets
- Centralized governance:** While it is beneficial for functions to own the data, it is important to strike a balance between decentralized data governance and centralization for the sake of interoperability. Complete decentralization should be avoided. To achieve this balance, a federated governance structure, which allows for self-governance and function-specific decision-making powers while adhering to a set of global rules that apply to all data products and their interfaces, is necessary. This approach ensures the creation of an interoperable ecosystem within the organization

By decentralizing responsibility and ownership of data, organizations can achieve greater agility. This allows enterprises to swiftly adapt to changing business requirements and explore new data products without disrupting the overall functioning of the organization. Additionally, data mesh contributes to enhancing data quality and reliability. With each team taking ownership of their respective data, they are accountable for ensuring its accuracy. Exhibit 4 illustrates the data mesh architecture.

EXHIBIT 4

Data mesh architecture

Source: Everest Group (2023)



It is important to note that data mesh is a relatively new concept and may lack the necessary implementation capabilities. Nevertheless, the growing demand from enterprises is transforming data mesh from a mere conceptual model into a thriving data ecosystem that actively supports its adoption.

Data vault – modernizing the design and architecture of data warehouses

Data vault is a modern architectural design for data warehouses that addresses the challenges posed by the storage and utilization of unstructured incoming data from various sources. Unlike traditional warehouse architectures store structured data, data vault stores data in its original, unaltered form, known as the as-is state. This approach is driven by the increasing demand for real-time data for decision-making and analysis purposes. The traditional approach of regenerating existing data based on predefined rules falls short in meeting these dynamic data requirements. However, data vault overcomes this limitation by providing flexibility in handling and integrating diverse data sources.

In the words of its inventor, Dan Linstedt, data vault can be defined as a detail-oriented, historical tracking system consisting of interconnected and normalized tables. It supports one or more functional areas of a business and combines the strengths of the third normal form (3NF) and star schema approaches. The design is flexible, scalable, consistent, and adaptable to the needs of modern enterprise data warehouses. It is specifically architected to meet the data management requirements of today's organizations.

Data vault improves flexibility and scalability in the data platform by eliminating the cleansing requirements typically associated with traditional star schemas used in traditional data warehouse modeling. Data vault consists of three essential components: hubs, which represent core business concepts; links, which depict relationships between hubs; and satellites, which store information about hubs and their relationships. Due to successful enterprise adoption, data vault modeling has evolved into data vault 2.0. This new iteration extends beyond the data warehouse and offers capabilities to handle cross-platform data persistence, multi-latency, multi-structured data, and massively parallel platforms.

Although data fabric, data mesh, and data vault differ conceptually, all three approaches have undergone initial development phases and have been explored. It is still uncertain which approach will garner more enterprise adoption and best suit the needs of modern enterprises. According to Everest Group's initial analysis, enterprises are at the early stages of their modern data platform journey and there is no single approach or architecture that dominates the field. In the next section of our report, we will examine the emerging trends in enterprise adoption of MDPs.

Enterprise adoption of MDPs

The data requirements of modern enterprises are driving data leaders to take strategic actions toward data modernization. While some enterprises have already embarked on this journey, many are still in the early stages. While one-off initiatives and process automation have been implemented in some organizations, they are no longer sufficient to stay relevant. It is vital for data and AI to be integrated into every decision at every level of the organizational hierarchy.

Another significant factor motivating enterprises to modernize their data strategies is rapid cloud adoption. Large enterprises across industries such as BFSI, HLS, and manufacturing, are shifting their

workloads to cloud. With critical workloads running in the cloud, there is a greater need to ingest, analyze, and contextualize data at a faster pace. This necessitates a shift from traditional data platforms to modern cloud-based platforms and architectures that offer scalability and flexibility for data-driven operations and decision-making.

The next-generation architectures we discussed earlier meet these enterprise requirements. While data fabric started getting adopted in early 2021, enterprises began embracing the data mesh approach in late 2022. Our research indicates that many large enterprises are actively incorporating these modern approaches into their data platforms to gain a competitive edge. Exhibit 5 outlines the adoption scale of these architectures across prominent verticals.

EXHIBIT 5

Enterprise adoption of modern data architectures

Source: Everest Group (2023)

Intensity: ● High ● Medium ● Low

Overall adoption of modern data management architectures across industries	Breakup across individual rating parameters		
	Active implementation	Talent readiness	Research and innovation
Healthcare and life sciences	● High	● High	● Low
Banking, financial services, and insurance	● High	● High	● High
Telecom	● High	● Medium	● Medium
Retail and consumer goods	● Low	● Low	● Medium
Media and entertainment	● Low	● Medium	● Low
Energy and utility	● Low	● Low	● Low

Note: The analysis is based on ~80 enterprises, selected as per their global fortune ranking, across six DAAI priority verticals. The companies are rated based on availability of case studies, presence of talent, job openings, and patent applications in the domain of data fabric, data mesh, and data vault

According to our analysis, roughly 33% of the analyzed enterprises have adopted at least one of the modern data management architectures. The adoption rate is particularly high in the Healthcare and Life Sciences (HLS) and Banking, Financial Services, and Insurance (BFSI) sectors. Following closely behind are the telecom, Retail and Consumer Goods (RCG), Media and Entertainment (M&E), and energy and utility sectors. Given the intricate nature of data requirements and the stringent regulatory pressures in these industries, BFSI enterprises are making substantial investments in developing talent capabilities specific to these evolving data management paradigms. HLS stands out with the highest number of job openings for roles related to these modern data architectures. While early adopters are already reaping the benefits of these advanced data architectures, others are making progress in scaling up their adoption efforts.

“ Unless I move everything into one central place, maintaining consistent standards for data management becomes impossible. However, with a mesh, all data adheres to the same standards without the need to relocate it. – Pawan Verma, Chief Data Officer, Industrial and Commercial Bank of China

Overall, our analysis indicates a promising market for these modern data management paradigms, with enterprises across various verticals expressing a positive outlook on future adoption. However, since these concepts are relatively new, organizations are encountering some challenges as they strive to advance and expand their adoption efforts.

Barriers that prevent businesses from developing a strong data platform

The increasing demand for AI is driving enterprises to adopt MDPs as a means to scale their AI initiatives. High-quality data is essential for any AI project, especially for the latest generative AI models. However, despite the eagerness to embrace MDPs, enterprises often face significant challenges that impede their progress. Below, we discuss some of these challenges in detail:

- **Inertia with the legacy setup:** Large enterprises often cling to their legacy data architecture despite being aware of data modernization technologies. This attachment stems from the substantial investments made in developing on-premise, monolithic architectures over the years. As technology becomes obsolete at a much faster rate, these legacy investments have turned into sunk costs for enterprises. What was once a strategic competitive advantage, modernization has now become an indispensable necessity. Without it, enterprises face a threat to their survival
- **Cultural mindset:** In today's business landscape, enterprises are striving to establish a stronger connection with their customers. To achieve this, it is vital to empower workers with data-driven intelligence, enabling them to make informed decisions and improve their productivity. Creating a strong data culture is essential for data leaders to effectively utilize data for their benefit. This culture should foster an environment where every employee, regardless of their position in the organizational hierarchy, feels comfortable with data and trusts it enough to make daily decisions. Unfortunately,

many large enterprises overlook the importance of instilling this culture and sense of data comfort among frontline employees

- **Lack of talent and the right skill set:** In the past, data management tasks were primarily handled by dedicated data management teams. However, as data platforms evolve and self-service capabilities become more prevalent, there is a growing demand to train business teams to handle complex data-related tasks. The hybrid work culture and expanding supply chains have led to vast and complex organizational structures. In a world where delivery personnel represent leading brands, it is vital to enhance the technical skills of employees throughout the organization
- **Problem of plenty:** The data platform space is dynamic and appealing, as evident by the increasing number of providers, which has surpassed 1000. This ecosystem continues to expand due to the attractive returns it offers. Particularly in modern data management paradigms like data fabric and data mesh, numerous participants are entering the market. However, with the growing ecosystem of platforms and tools, enterprises encounter challenges in effectively navigating and selecting the most suitable options for their needs

Best practices for building an MDP

Each organization has its own distinct level of data maturity, making it clear that there is no universal approach to constructing an MDP. To navigate the complexities involved, enterprises must take into account several best practices when embarking on their MDP development journeys.

- **Not just about the tool:** Data modernization encompasses more than simply choosing the appropriate tools to build a robust platform. It is a comprehensive journey that entails multiple steps and challenges. It goes beyond the technical readiness of enterprises to adopt modern data tools; organizations must also be mentally prepared to embrace a data-driven culture. Data modernization is not a one-time endeavor but an ongoing process that necessitates constant monitoring and evaluation to ensure that the updated infrastructure aligns with the evolving business needs. This entails ongoing maintenance and upgrades to the technology stack, along with continuous training and development of employees to stay abreast of new advances and best practices in data management
- **Big bang vs. a step-by-step approach:** There are two approaches that enterprises can consider for data modernization: the big bang approach and the step-by-step approach. The big bang approach involves a comprehensive overhaul of an organization's data infrastructure and technology all at once, while the step-by-step approach takes a more incremental and phased approach to data modernization. Before selecting the appropriate approach, enterprises should carefully evaluate their current data maturity, available resources, and risk appetite. It is important to avoid developing a herd mentality and instead make a well-informed decision. Incorrect assessment may result in enterprises over or under strategizing their data modernization journeys, leading to wasted resources and ineffective outcomes
- **Value realization and cost consciousness:** Most modern data management tools often highlight their lean pricing structures and how they can assist enterprises lower their IT spending through cloud deployments. However, enterprises need to exercise caution and note that these pricing structures can sometimes be misleading, as infrastructure costs may be bundled with data costs. Additionally, despite the widespread adoption of cloud-based data platforms, businesses have yet to fully

experience the cost advantages that were promised. As a result, enterprises are now conducting careful evaluations of their cloud-only strategies and are opting for hybrid architectures that can better support their business-critical data requirements

- **Having the right data tool is important but having the right talent is nonnegotiable:** Modern data platforms are inherently complex and demand specialized skills and knowledge to operate effectively. Therefore, to fully leverage the capabilities of advanced data platforms, businesses must invest in recruiting and retaining the right talent. It is evident that the innovation cycles for breakthrough technologies have dropped significantly. As a result, hiring should no longer be considered as a one-time activity but rather an ongoing process. Enterprise leaders must prioritize continuous training and development initiatives for their teams to ensure they stay up-to-date with the latest data tools and technologies
- **Selecting the right-fit provider:** As the interest from enterprises continues to grow, the data platform space becomes increasingly crowded with numerous providers entering the market. This saturation makes it challenging for enterprises to navigate and identify the best choice for their data stack. In a modern enterprise, effective collaboration between data scientists, analysts, developers, and machine learning creators is vital. Therefore, leaders in data management must exercise caution and thoroughly evaluate providers to ensure they can meet the diverse requirements of all stakeholders

While there is no definitive approach or comprehensive playbook for data modernization, the best practices mentioned above significantly enhance the chances of success for enterprises. These best practices serve as a foundation for enterprises to initiate their modernization journey and create a roadmap for implementation. It is important to recognize that data modernization is an ongoing process, and as such, enterprises should adopt a long-term vision and strategy and avoid temporary, quick-fix initiatives that only address immediate issues, as they may not align with the organization's long-term goals.

Creating a strong foundation for data-driven innovation: a case study on TUI's MDP journey with Sonata Software

Company background, business objectives, and challenges

TUI, a Germany-based leisure, travel, and tourism enterprise, encountered several challenges in managing and analyzing their data. It recognized the need to transition from its current warehouse-based approach to a cloud-based solution. Its traditional on-premise data management systems lacked scalability and struggled to handle the increasing volume and variety of data sources. Additionally, TUI sought to incorporate data mesh components to modernize its data platform. To address these challenges and modernize its data architecture, TUI partnered with Sonata Software.

Solution

TUI initiated the transformation of its warehouse-based data infrastructure by implementing a cloud-based data lakes solution. This solution empowered them to store raw and unstructured data, establishing the groundwork for advanced analytics and insights. Additionally, as the volume of crucial data components stored in TUI's on-premise warehouses increased, TUI migrated to Snowflake's cloud-based warehouse systems, further enhancing scalability and flexibility. To accommodate the evolving requirements of TUI's data architecture, they introduced event-driven capabilities that facilitated responsiveness to events and efficient event-driven processing. This enabled the company to adapt and handle big events alongside big data workloads.

As TUI progressed along the data maturity curve, it incorporated several critical elements of a modern data platform. These included data governance, metadata management, pre-built visualizations, and self-monitoring features. These additions aimed to assist TUI's data teams in comprehending the data flow and identifying cost centers. While TUI already possesses a strong centralized data management system, it faces a significant challenge. Despite the businesses' expertise in their respective domains, the central team is responsible for managing the data. To overcome this challenge, TUI is now integrating components of data mesh and distributed domain ownership. This approach ensures that individuals who understand the data take ownership of it and deliver it as a product to other parts of the business.

Outcome

Sonata Software played a pivotal role in TUI's MDP journey. While the data mesh implementation at TUI is still in progress, the company has already begun reaping the rewards of scalability and data-driven decision-making. This has established a solid data foundation for future growth and innovation. As TUI continues to advance its data estates, it strives to explore additional opportunities in cost control and implement advanced natural language capabilities to support the growing the customer demand for natural language processing-based interfaces.

Future outlook

In this report, we explored the urgent requirement for enterprises to reimagine their traditional data platforms and initiate their data transformation at the earliest. In an era characterized by democratic intelligence and scaled AI, possessing an MDP becomes essential for enterprises. As we progress in this dynamic landscape, we see three prominent trends that are shaping the future of MDPs.

Convergence of data mesh and data fabric use cases

While various data management architectures offer their own advantages, two notable architectures, data mesh and data fabric, have distinct perspectives on centrality. Data Fabric promotes a centralized view of data management, whereas data mesh advocates for a decentralized approach, with business domains taking ownership of their respective data. An MDP will incorporate elements and best practices from both approaches, striking the right balance between centralization and decentralization of data. Industries with complex organizational structures, spanning multiple geographies, and encompassing multiple business units, such as BFSI, telecom, and healthcare, stand to gain significant benefits from this amalgamation of centralized and decentralized elements in their data platforms. In such cases, both data mesh and data fabric can coexist, providing a powerful framework for data management.

More unicorns and rising Mergers and Acquisitions (M&A) activities

The surge in digital activities and proliferated data generation have already drawn significant attention to data platforms among enterprises. However, the recent hype around generative AI has emphasized the importance of accurate and high-quality data for the optimal functioning of next-generation AI systems. This trend highlights the increasing volume of data that will be generated by generative AI models in the future, further solidifying the necessity of robust data platforms. Clearly, this is a lucrative domain to be involved in, and going forward, we can expect to see more unicorns in this field. Additionally, there will likely be a rise in M&A activities as organizations seek to capitalize on the opportunities presented by this evolving space.

The increasing significance of cloud and data economics will drive the demand for data observability in the MDP

The ongoing financial crisis has underscored the significant importance for enterprises to rethink their budgets. In this context, enterprises are increasingly reevaluating their infrastructure and data costs, and they are displaying a strong interest in data observability tools to monitor their spending areas. By providing continuous oversight of storage and identifying stale, corrupted, and unused data, data observability facilitates the optimization of cloud resources. It also contributes to lowering operational costs by proactively preventing the processing of corrupt data before it enters the system, thereby saving on computation costs. Owing to its widespread benefits, we see a rich ecosystem of participants emerging in this domain and anticipate substantial growth in the market.

Everest Group is a leading research firm helping business leaders make confident decisions. We guide clients through today's market challenges and strengthen their strategies by applying contextualized problem-solving to their unique situations. This drives maximized operational and financial performance and transformative experiences. Our deep expertise and tenacious research focused on technology, business processes, and engineering through the lenses of talent, sustainability, and sourcing delivers precise and action-oriented guidance. Find further details and in-depth content at www.everestgrp.com.

This study was funded, in part,
by [Sonata Software](#)



**For more information about
Everest Group, please contact:**

+1-214-451-3000

info@everestgrp.com



**For more information about this
topic please contact the author(s):**

Vishal Gupta, Vice President

vishal.gupta@everestgrp.com

Priya Bhalla, Practice Director

priya.bhalla@everestgrp.com

NOTICE AND DISCLAIMERS

IMPORTANT INFORMATION. PLEASE REVIEW THIS NOTICE CAREFULLY AND IN ITS ENTIRETY. THROUGH YOUR ACCESS, YOU AGREE TO EVEREST GROUP'S TERMS OF USE.

Everest Group's Terms of Use, available at www.everestgrp.com/terms-of-use/, is hereby incorporated by reference as if fully reproduced herein. Parts of these terms are pasted below for convenience; please refer to the link above for the full version of the Terms of Use.

Everest Group is not registered as an investment adviser or research analyst with the U.S. Securities and Exchange Commission, the Financial Industry Regulatory Authority (FINRA), or any state or foreign securities regulatory authority. For the avoidance of doubt, Everest Group is not providing any advice concerning securities as defined by the law or any regulatory entity or an analysis of equity securities as defined by the law or any regulatory entity.

All Everest Group Products and/or Services are for informational purposes only and are provided "as is" without any warranty of any kind. You understand and expressly agree that you assume the entire risk as to your use and any reliance upon any Product or Service. Everest Group is not a legal, tax, financial, or investment advisor, and nothing provided by Everest Group is legal, tax, financial, or investment advice. Nothing Everest Group provides is an offer to sell or a solicitation of an offer to purchase any securities or instruments from any entity. Nothing from Everest Group may be used or relied upon in evaluating the merits of any investment.

Do not base any investment decisions, in whole or part, on anything provided by Everest Group.

Products and/or Services represent research opinions or viewpoints, not representations or statements of fact. Accessing, using, or receiving a grant of access to an Everest Group Product and/or Service does not constitute any recommendation by Everest Group that recipient (1) take any action or refrain from taking any action or (2) enter into a particular transaction. Nothing from Everest Group will be relied upon or interpreted as a promise or representation as to past, present, or future performance of a business or a market.

The information contained in any Everest Group Product and/or Service is as of the date prepared, and Everest Group has no duty or obligation to update or revise the information or documentation. Everest Group may have obtained information that appears in its Products and/or Services from the parties mentioned therein, public sources, or third-party sources, including information related to financials, estimates, and/or forecasts. Everest Group has not audited such information and assumes no responsibility for independently verifying such information as Everest Group has relied on such information being complete and accurate in all respects. Note, companies mentioned in Products and/or Services may be customers of Everest Group or have interacted with Everest Group in some other way, including, without limitation, participating in Everest Group research activities.

About **Sonata Software**

In today's rapidly changing world, the modernization of technology and business processes is no longer a luxury but a necessity for organizations that aspire to stay competitive.

Sonata is a leading Modernization Engineering company, powered by its unique Platformation™ framework that brings together industry expertise, platform technology excellence, design innovation and strategic engagement models to deliver sustained value to customers.

We partner with Fortune 500 clients to help them reimagine their business processes as part of modernization programs. As part of this phase, we drive business enablement sessions, consulting, and business processes to be aligned with best-in-class industry-specific practices.

Our outcome-based Modernization Services include Cloud, Data, Dynamics, Managed Services, Automation and Digital Contact Center. We offer services and solutions around newer technologies like Generative AI, MS Fabric, and other modernization platforms.


Sonata has a strong global presence in key regions including US, UK, Europe, APAC, and ANZ. We are a trusted partner of world leading companies in TMT, Retail & CPG, Manufacturing, BFSI and HLS space.


Our partner ecosystem boasts of strong relationships with Microsoft, AWS, Salesforce, Snowflake, Google, and Industry partners and is critical to the value we create with our clients.


Sonata is committed to the Environmental, Social, and Governance (ESG) charter. We recognize that our success is tied to the well-being of our stakeholders, including our employees, clients, shareholders, and the communities in which we operate.




www.sonata-software.com

 www.linkedin.com/company/sonata-software

 www.facebook.com/sonatasoftware

 https://twitter.com/Sonata_Software

 www.instagram.com/sonata.software