WHITE PAPER

Beyond Technology:

Creating Business Value with Data Mesh
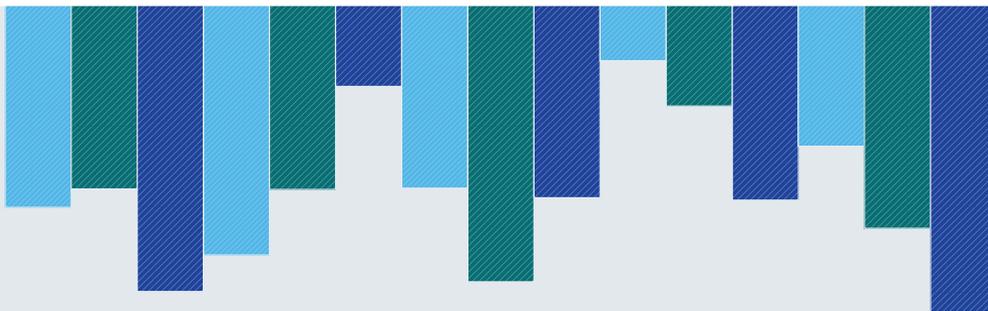
There has been tremendous advancement in the data and analytics industry over the past two decades. Growth in available data, both in terms of volume and in diversity of sources, has been exponential and those businesses that are able to effectively utilize this data, using advanced analytics and artificial intelligence (AI), have unlocked a significant advantage over their competitors.

However, the organizations that actually experience these benefits remain in the minority. For the rest, the promise of AI and data science remains elusive, and even effective operational reporting can be a challenge. Despite continued investments in data management initiatives and the latest technologies, these organizations continue to struggle with the first step: enabling their employees to find and consume the data they need to make better decisions and successfully deliver value.

This struggle with the first step leads us to an important question: Why does providing secure and reliable access to high-quality data remain so difficult?

The surprising answer is that, for the most part, technology is not the core issue. Instead, most companies find that the operational and organizational issues they face are their greatest impediments. Overcoming these challenges requires the business to change how it perceives the value of data, how business stakeholders view their responsibilities for effective data stewardship, and how data management teams operate to support their business.

Data mesh is a new approach to creating data solutions that seeks to address these challenges by applying modern organizational and engineering practices. By adopting the principles and practices of data mesh, IT organizations are able to respond more quickly to the evolving data needs of their organization and collaborate more effectively with business partners to create insights.

We have sponsored this research by Harvard Business Review Analytic Services to share the experiences of business and technology leaders as they seek to deliver strategic insight to optimize their business and provide better experiences for their customers. We explore the challenges they face in creating value from their data and the impact of adopting the data mesh paradigm to address these challenges.

In our digital economy, the ability to produce, share, and act on business intelligence has never been more critical to an organization's success. Read on to discover the key strategies for winning with data mesh.

John Spens
Director, Data & AI
Thoughtworks

Beyond Technology:
Creating Business Value with Data Mesh

Turning data into actionable insights for better business outcomes has been the goal of organizations for well over a decade now. The pandemic has further underscored the need for data-driven transformation, just as it has revealed the importance of having visibility into supply chains. Companies can benefit exponentially from using data to more precisely understand their customers' and employees' needs and expectations.

Businesses are making the requisite investments in data analytics technologies and skill sets to remain competitive and boost their growth. Fifty-five percent of IT decision makers plan to increase investment in data-focused budgets over the next 12 to 18 months, according to a 2022 study by Foundry.[1] The same study found that these companies will spend an average of $12.3 million in the coming year on data-driven initiatives, including software, services, training, consulting, and other related costs.

Companies' success with data analytics will depend on how they approach their transformations. For many, IT is the warden of data, but that isn't always the most effective path to becoming a wholly data-driven enterprise. "Many data analytics initiatives do not yield satisfactory business outcomes, because with traditional data architecture, centralized IT teams act as gatekeepers of data," says Daniel J. Abadi, professor of computer science at the College Park, Md.-based University of Maryland. "They are experts at data management tools but not the data itself." In this centralized data system, technology is considered a goal and not an enabler to achieve business results—it focuses on the output of data-driven initiatives and not their outcomes.

In recent years, organizations have been turning to a new approach to data management called "data mesh." Data mesh is a federated approach to data management, unlike more traditional centralized approaches where data

HIGHLIGHTS

For an organization to become data-driven and empower business teams with actionable insights, data needs to become a product, essentially a data set that is consumable by anyone in the organization.

Data mesh is a federated approach to data management, unlike more traditional centralized approaches where data sits in warehouses or lakes and is controlled by a central IT department.

The success of data mesh depends on the organization's ability to engage technology and business executives and change the organization's social infrastructure.

"Many data analytics initiatives do not yield satisfactory business outcomes, because with traditional data architecture, centralized IT teams act as gatekeepers of data," says Daniel J. Abadi, professor of computer science at the University of Maryland.

sits in warehouses or lakes and is controlled by a central IT department. With data mesh, the control and ownership of data are accessible to different domains throughout the organization, and these domains are responsible for handling data and following organization-wide governance rules. With this approach, data is handled by teams in functions besides IT, closer to the business, so the data is better suited to business needs.

The essence of data mesh is that it puts the customer first and focuses on data yielding outcomes instead of outputs. Just like with any other product in the marketplace, the value of the data product, or data set, depends on whether data consumers need and use it. This federated approach aims for different outcomes and follows various processes and workflows as opposed to traditional centralized approaches; thus, it requires the creation of new roles, teams, responsibilities, and accountability. The biggest impact of data mesh is on the responsibility for data, which resides with business domains rather than with IT.

The success of data mesh depends on the organization's ability to engage technology and business executives and change its own social infrastructure. "My recommendation is to approach data mesh with the 70-20-10 mindset. Seventy percent of the focus should be on change management, 20% on data and technology, and 10% on models and AI [artificial intelligence]," says Sanjeevan Bala, group chief data and AI officer at ITV, a London-based free-to-air television network.

This report discusses the benefits of taking a decentralized approach to data and the disadvantages of a centralized data management system. It explores how companies are implementing data mesh, what challenges they are facing, and how they are overcoming them. The report also focuses on how to form data mesh teams and how best to create and track metrics.

An Embarrassment of Riches

Data is undoubtedly a strategic asset. It can help answer critical questions that decision makers have about how best to grow their organizations, including which products customers are buying and how satisfied they are with those products, and whether employees are productive and engaged. Data also can answer questions about the reliability of supply chains and what demand will look like in the future.

Insights from data have the potential to boost companies' competitive advantage and spur growth by improving internal business processes (50%), customer insights and engagement (46%), and customer service (43%), according to the Foundry study. **FIGURE 1**

But for many organizations, data is more an embarrassment of riches than a strategic asset because they are not able to turn data into insights and derive optimal business value from it. Data volume is increasing by 63% per month at large organizations (those with at least 1,000 employees), and the mean number of data sources per organization is 400, with one-fifth of organizations drawing from 1,000 or more data sources.[2] These data sources can be structured or unstructured; internal or external; raw or formatted; and

FIGURE 1

Looking for Business Insights

Top business goals driving investments for data-driven initiatives

50%	Improve/automate internal business processes
46	Improve customer insights and engagement
43	Improve customer service/support
43	Improve/automate IT operations
36	Improve existing products
36	Improve information security/cybersecurity
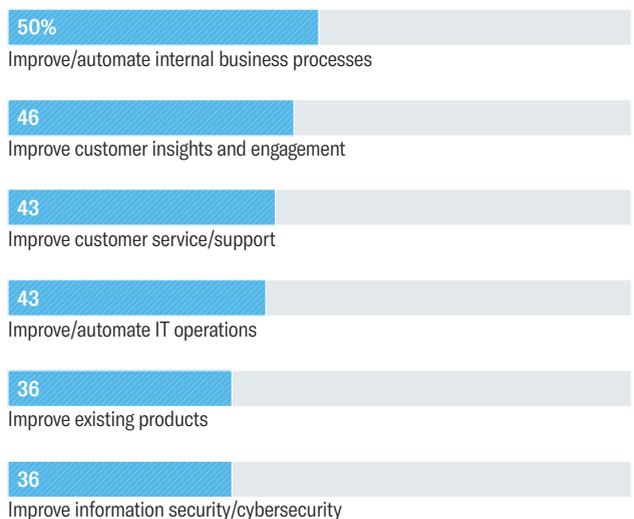
Source: Foundry, 2022

visual, transactional, or collected by sensors embedded in machines, among other types.

The amount of data does not help if the data itself is not trustworthy. "Without focus on data governance, there's an overall mistrust in data because even if it is of good quality, it can be used in the wrong context and under wrong assumptions," says George Firican, director of data governance and business intelligence at the University of British Columbia in Vancouver. "A lot of organizations are aiming to become data-driven and are investing more and more in data analytics, but they often miss on investing in a strong foundation created by data governance."

Even those organizations that focus on data governance—or a set of internal data standards meant to ensure that data is consistent and trustworthy—still may not be able to use data for business decision making due to traditional, centralized approaches to data management. Typically, data is sitting in warehouses (repositories for structured, filtered data that already has been processed for a specific purpose) or data lakes (storage repositories that hold a vast amount of raw data in its native format until it is needed for analytics applications), both of which are typically controlled by IT. Potential consumers, be they business leaders or data analysts, may not be fully aware of what all the data is, how to access and understand it, and how it can be useful for them to grow the business.

For bank traders, for example, going through the central data group to get insights from price fluctuations defeats the purpose because it takes too long, says Sheetal Pratik, director of engineering and data integration at Adidas and formerly head of data integration at Saxo Bank, a Copenhagen, Denmark-based investment bank specializing in online trading and investment, where she rolled out a data governance self-service platform.

In a centralized system, those in IT have the responsibility of transforming data to produce relevant insights without being domain experts. They need to rely on subject matter experts to work together on the rules about how to transform and govern the data. And things don't always work the first time when processing and analyzing data or perfecting the algorithm; the programmers need to go back to the subject matter experts to make adjustments, says Pratik. This requirement creates a lot of back-and-forth among different groups working separately because each group has different skill sets. Additional delays and bottlenecks occur because the central IT group is handling requests from the whole organization, and requests for the analysis of different data sets from multiple domains sit in a queue and wait their turn.

Another drawback of a centralized system is that it can endanger the reputation of data. Production of data that is not trustworthy or relevant is more likely to occur in a centralized system because centralized data management

"Data doesn't deliver value up to its full potential due to traditional practices of centralization of enterprise data in warehouses or data lakes," says Omar Khawaja, global head of business intelligence and analytics solution and services teams for Roche Diagnostics.

teams tend to be more disengaged from the business side, and thus they have less understanding of what the data means. Once confronted with data that is untrustworthy, business executives may find it hard to regain confidence in using data. That point is true even if untrustworthy data appears in models or demonstrations during the testing stage.

Functions across the business require access to self-service data tools. Bank traders, for instance, need them to generate useful data-driven insights in real time. Such insights can only be produced by people who understand the bankers' trading needs. Marketers, too, need self-service data tools to easily look up insights from customer data to determine what advertising content to distribute. And planners in a manufacturing plant who decide on inventory levels or material purchases need such tools to access timely production and sales data.

It's no wonder, then, that 83% of IT decision makers agree that providing self-service tools to make data more accessible to business is a top priority, according to the Foundry study.[3]

But that's not how the generation of data-driven insights works in most of today's organizations. Nearly a third of organizations say that data analytics solutions are available for all or most users, but to access and use the data, they need to have specialized data analytics or computer science skills. Another 41% or so of organizations say that data and analytics solutions can only be used by a few skilled teams, such as analytics teams and IT departments. Just one in five (21%) have data and analytics solutions that are available and easy for all users to leverage.[4]

"Data doesn't deliver value up to its full potential due to traditional practices of centralization of enterprise data in warehouses or data lakes," says Omar Khawaja, global head of business intelligence and analytics solution and services teams for Roche Diagnostics, a Basel, Switzerland-based multinational health care company. "How can a centralized team of 50 or 100 people in an enterprise which has 100,000

> "My recommendation is to approach data mesh with the 70-20-10 mindset. Seventy percent of the focus should be on change management, 20% on data and technology, and 10% on models and AI," says Sanjeevan Bala, group chief data and AI officer at ITV.

people make the whole company data-driven? That's just a fool's paradise."

Empowering People with Data Mesh

For an organization to become data-driven and empower business teams with actionable insights, data needs to become a product, essentially a data set that is consumable by anyone in the organization. Just like any other successful product, a data product must create and fulfill a need or meet data consumers' expectations, be easily accessible, and be timely.

"Data should be accessible, interoperable, and reusable," says Khawaja. "This means that the data product is published on an internal data marketplace so that people like data scientists, business analysts, or teams sitting in other domains can access the data. They can then either generate insights on their own or use those data points to create other data products to generate further insights by connecting the dots between different domains."

Data mesh offers a methodology for creating data products. Adidas' Pratik, who implemented data mesh when she was at Saxo Bank, sees its value as giving control of the data back to the business domains while setting up organization-wide protocols.

With this approach, business units become accountable and responsible both for how they produce and use the data for their own business and for how useful their data products are to the whole organization. The way to implement data mesh, says Pratik, is to enable the business teams with technology by combining business and technical teams. This way, the integrated team possesses all the required knowledge and skill sets to create a data product from end to end without having to rely on handoffs to other teams.

ITV's Bala, who started pursuing a data mesh approach two years ago, explains, "Rather than creating a centralized team, a sort of an ivory tower, we fully decentralized both technology and data teams in such a way that they could sit with the business units, in effect going native into the business."

ITV has already seen multiple benefits from adopting data mesh. For example, data sharing and collaboration between the sales and content creation teams are bearing fruit in contextual advertising. The teams that create TV shows made a lot of metadata available, which helped the sales team better understand what shows are about. When a certain product is spoken about in positive terms on a show, that helps the sales team sell ads for that product. So, if there is talk about coffee on an episode of *Love Island*, a reality dating show, the coffee advertising does really well. "It's a natural fit because there is no clash with advertising," says Bala. "Data mesh opened up collaborations that didn't exist before."

Speed is among the top benefits of data mesh. With the data products team in control of producing and using data, there is no dependency on any cross-organizational team, which eliminates backlogs. Users are able to consume data on a real-time basis. For ITV, having relevant data products sped up the ability to activate segments promoting shows to targeted viewers from three months to three minutes. "We see data mesh as an opportunity to leapfrog and accelerate our transformation," says Bala.

The New Value-Driven Social Architecture

Data mesh disrupts the traditional ideas of a centralized system and the foundations of data management, which means the approach must be embraced culturally. "It's a big paradigm shift in terms of the way of working, the mindset, team structure, technology, and architecture," says Roche's Khawaja.

He explains all the major elements that were involved in the shift to data mesh at Roche Diagnostics. "Instead of funding projects, we funded teams. Instead of having only IT people involved in data analytics, we had cross-functional teams," Khawaja asserts. "Instead of doing projects, we are talking about product thinking. Instead of working on a classic on-premises, non-scalable infrastructure, we are using our state-of-the-art, self-service data infrastructure. Instead of just creating dashboard after dashboard, we are talking about creating reusable data products."

The product teams are responsible for the life cycle of that data product from end to end, adds Khawaja. With the implementation of data mesh, the silos between IT and other business functions disappeared, and people joined product teams from different domains. "It no longer matters where people come from as long as they follow certain data governance boundaries," he notes.

In effect, data mesh helps solve the perennial challenge of trying to make businesspeople understand technology teams, and vice versa. Once people with different skills collaborate in a data product team, they start complementing each other. For example, a data product manager can talk to a colleague in IT about the availability of certain types of data, and then they can discuss the potential algorithmic uses of this data with data analysts and teammates who are running the business. This approach makes the team aware of what data they have and how to use it to create business value even before they start building the data product, explains Kevin Hong, professor of business technology at the University of Miami's Miami Herbert Business School.

"An important advantage of data mesh is that data scientists embedded in the domain teams are more likely to ask the right questions of what can be accomplished with data because they work closely with business leaders," says Hong. "[Within a centralized approach,] data analysts often know how to analyze data, but they may not know why they are analyzing it."

But while data mesh solves many problems, this approach also alters people's roles and responsibilities, redefines individual goals and accountability, and goes across traditional reporting structures. Hong, who has taught multiple executive courses on data analytics, says that executives often point to power plays in organizations as the reason why data initiatives do not translate to business outcomes. Simply put, data gatekeepers who under a centralized system have power over the data—with everybody coming to them with requests—lose that power when control of data is moved to the domains.

"Organizations are doing a lot of things in terms of adopting the latest-generation technologies to enable data mesh. However, where I've typically seen major problems is [not] educating people about the new roles and the accountability that comes with it," says Pratik. She suggests holding training workshops and educating both the business and the technology sides about new ways of collaborating.

"Communication [about new roles and responsibilities with data mesh] is key and often overlooked," adds the University of British Columbia's Firican. "The necessary policies, processes, standards, roles, and responsibilities need to be established in order to clarify who does what, when, and how, but that's not enough. Communication and overall change management need to be tightly coupled with data governance."

The goal of data mesh is to drive not technical outcomes but business outcomes. That's why it makes sense for organizations to start any data mesh transformation by establishing business value that the data mesh will generate and engaging business executives in the process from the very beginning. While many business executives may be technology savvy, the idea of a shift to data-driven management with data mesh is still

> "[Within a centralized approach], data analysts often know how to analyze data, but they may not know why they are analyzing it," says Kevin Hong, professor of business technology at the University of Miami's Miami Herbert Business School.

forcing a change in the way they have been operating, often very successfully, for years.

The first step that Bala took when implementing data mesh at ITV was to look at the value chain. He engaged business stakeholders by asking how using data and artificial intelligence might change what they do when they produce, promote, distribute, or monetize content. "We were helping stakeholders reimagine how their part of the business could operate differently using data and AI," says Bala.

The paradox of data mesh implementation is that while it ultimately drives business value, it is most often led by technology executives who do not control marketing, sales, or production. It can be overwhelming for technology executives to convince business leaders to embrace data mesh, says Pratik, and it's common for a tech-led initiative to reach a dead end at the business level. "If [the data mesh transformation] becomes business-led, it will become magical," she says.

For data mesh to succeed, the C-suite should be involved. The data mesh transformation at ITV was part of the company's innovation agenda to drive growth, and it was supported by the CFO, the management board, and the board of directors. Since it was viewed as an organizational and cultural change, versus just a technology implementation, the human resources department also played a big role.

To get the business owners to embrace data mesh in their domains, Bala got the right stakeholders involved in mapping the business benefits that could be achieved with data mesh. He used a "value tree," which showed which data products lead to the organization's desired business outcomes. For example, a data product that could help predict how many and which viewers are likely to churn and what the churn rate depends on would help inform the business side about what changes to make to retain viewers and ultimately protect revenue.

ITV runs control groups against the organization's data product experimentation to prove value generated by

"The biggest challenge to succeeding with data mesh is creating an internal data marketplace with data producers and data consumers," says Sandra Cannon, the first chief data officer at the University of Rochester and a data management consultant.

using data products. Control groups allow the company to understand which data products are working and which data-driven KPIs lead to desired business outcomes. "The business knows what lever to pull to affect internal KPIs—such as monthly active user target or viewing hours targets or churn—and how these internal data-driven KPIs impact revenues," says Bala. "That's how we got the business on our side." Bala's team now finds itself in the enviable position of being asked by other business executives for help with a data mesh approach to growing their business lines.

Other companies also have had significant success with data mesh. At Roche, data used by manufacturing planning teams is now available every day, instead of every month or every quarter. Having implemented data mesh and thus eliminated the bottlenecks of a centralized system, the organization is now able to produce actionable insights from data much faster. Data products highlight up-to-the-moment details of the different materials required in the inventory. And access to timely data insights leads to savings in maintaining material inventory and improved efficiency in the manufacturing process. Just last year, a reduction in the stock safety requirements, achieved thanks to having up-to-date stock data, delivered savings in terms of capital investment.

Creating an Effective Data Products Marketplace

"The biggest challenge to succeeding with data mesh is creating an internal data marketplace with data producers and data consumers," says Sandra Cannon, the first chief data officer at the Rochester, N.Y.-based University of Rochester and now a data management consultant to government and not-for-profit organizations. "You need to make sure that you've got the right incentives built into the roles and responsibilities for the people who are doing all the work. You need to address the issue of what's in it for them."

Much of the work that goes into creating data products is related to data governance. "Data governance is the biggest stumbling block of data mesh. If you don't do it right, data mesh is not going to work," says the University of Maryland's Abadi, who performs research on database system architecture and implementation.

With data mesh, data management standards are distributed, which means that every domain working on data products needs to follow them. The first step is to create standards that are distributable and applicable across the organization. This step requires creating a master data management standard that describes standard definitions and rules—a common data language, so to speak—to make sure things don't get lost in translation between the domains, explains Abadi.

The next step is making sure that data teams are maintaining these standards in their products. "Creating consistent data requires a lot of work from upstream data products' producers," says Abadi. When, for instance, an American and a European company merge, data products involving the newly combined company's prices need to be compatible. However, the different currencies might mean that data sets cannot be automatically combined. The pricing then needs to be accurately converted and cleaned up so that when data consumers query different data products, including pricing, they get correct information. And different functions within both companies may require different data points. A finance department of the overall organization may need converted pricing to project overall sales, but marketing departments in different countries may need pricing in local currencies, with or without discounts, to establish how to tailor loyalty programs to their customers.

In order to incentivize and reward data producers, it's necessary to track all the producers' contributions to the data products, notes Abadi. Since many of the products incorporate other data elements and are repackaged as new products, the majority of the data teams are both data consumers and data producers. They consume the upstream data products, and they produce new ones, which are consumable downstream. As an example, the upstream data team may track transactions, while a downstream team may take this data product and add other data—such as customer locations or identities—and make it more valuable.

"If you create new data products with references to the original data sets, you know where the data is coming from, and if the product is valuable, you know whom to reward," says Abadi. "Then the producers are going to be incentivized and create more products, which will make other people's lives easier."

"The business knows what lever to pull to affect internal KPIs—such as monthly active user target or viewing hours targets or churn—and how these internal data-driven KPIs impact revenues. That's how we got the business on our side."

Sanjeevan Bala, group chief data and AI officer at ITV.

> "Technology is a key enabler, but without the focus on people, technology-driven initiatives will become derailed," says Roche Diagnostics' Khawaja.

Spearheading Innovation with Data Mesh

Data mesh has the potential to help companies succeed at digital transformations aimed at making them data-driven because it addresses the need to transform social architecture while implementing the technology architecture. But data mesh can bring challenges that organizations need to be ready to address.

Traditionally, companies have been focusing on technology while not paying enough attention to change management, which is why so many digital transformations have failed. "While the right technology choices will enable teams to work in data mesh, do not start with technology," says Khawaja. "Technology is a key enabler, but without the focus on people, technology-driven initiatives will become derailed."

Bala advises starting the implementation of data mesh with the business value and focusing on the "last mile," which is when data consumers get the data product "in their hands." This moment of truth reveals whether the created data products are useful. Before starting to build any models, organizations need to make sure that marketers, operators, planners, and other business lines will have easy, self-service tools to access the models. They also need to be involved in the process to make sure the models will yield business outcomes.

It is important that organizations design a functioning data products marketplace. An effective marketplace can be achieved through the promotion of a healthy data producers and consumers relationship based on a system of incentives in which everyone who produces value gets rewarded. Organizations should set up the incentives, provide self-service technology platforms, and promote teams' collaboration in a way that enables more and more people to become value-add data producers.

Adopt a "thin-slice, iterative approach," says Bala, which means working with small, manageable pieces of infrastructure and small amounts of data and developing these thin-slice use cases end to end. After showing value from these first use cases, organizations can build from there. While the start of data mesh may be based on a thin-slice strategy, Bala sees the potential of data mesh not just going enterprise-wide but also crossing the boundaries between companies and industries.

"As different organizations enable data mesh, it will allow us to use data across industries. With the hyperconnectivity going up a level and crossing sectors, we will be able to spearhead huge amounts of innovation," he says.

Endnotes

1   Foundry, "Data & Analytics Study," 2022.

2   IDG Market Pulse Survey, "Optimizing Business Analytics by Transforming Data in the Cloud," 2022.

3   Foundry, "Data & Analytics Study," 2022.

4   Ibid.

Harvard Business Review

ANALYTIC SERVICES

MCCRE32170922