



Practical Approaches to Implementing a Data Mesh

BY JAY PISCIONERI
MAY 2023

This publication may not be reproduced or distributed
without Eckerson Group's prior permission.

About the Author



Jay Piscioneri has over 25 years of experience in data technologies including data warehousing, business intelligence, data quality, and data governance. He's worked with organizations in a wide variety of industries helping them plan and build their next generation data platforms. As a leader of many initiatives, Jay brings extensive experience in the nuts-and-bolts of implementing data solutions, as well as the organizational challenges that go with adopting new priorities, processes, and tools.

About Eckerson Group

Eckerson Group is a global research, consulting, and advisory firm that helps organizations get more value from data. Our experts think critically, write clearly, and present persuasively about data analytics. They specialize in data strategy, data architecture, self-service analytics, master data management, data governance, and data science. Organizations rely on us to demystify data and analytics and develop business-driven strategies that harness the power of data. [Learn what Eckerson Group can do for you!](#)



About This Report

To conduct research for this report, Eckerson Group interviewed data mesh practitioners from both user and vendor organizations. This report is sponsored by Cinchy which has exclusive permission to syndicate its content.

Table of Contents

Executive Summary	4
Understanding Data Mesh	6
Best Practices	12
Conclusion	21
About Eckerson Group	22

Executive Summary

Data mesh offers a new paradigm for fulfilling the promised value of data. It rejects long-standing centralized data architectures such as the data lake and the data warehouse and their associated centralized teams. Instead, it decentralizes both data ownership and the data itself, shifting them toward the functional domains that create and use data to run their business.

Data mesh is designed for the highly distributed and diverse world of modern data environments. Its four pillars (domain ownership, data as a product, the self-serve data platform, and federated computational governance) are the foundation for a very different approach to managing data for analytics.

Many organizations find this approach compelling. They recognize the flexibility and delivery advantages of decentralized domain ownership and data as a product. They see how a self-serve data platform removes overhead and friction for both developers and consumers of data. And they want to mature and expand their data governance practices without slowing down development work or stifling innovation.

However, practitioners implementing data mesh must work within the bounds of what their organization's culture, existing technologies, and resources will allow. This report explains how data mesh is a journey that unfolds iteratively according to a company's needs and readiness for the changes that it imposes.

Key Takeaways

- > Data mesh distributes data ownership from a central data team and central data systems to business domain teams who create and use their data.*
- > Data mesh is not for every organization. You need sufficient data size and complexity to justify the investment in a data mesh program.*
- > Domain ownership and data as a product require a change in mindset. Engineering teams must expand their focus to include developing data solutions for their stakeholders, which is likely to be a significant change in how they work.*
- > Product thinking applied to data is critical to data mesh. Domain teams must discover stakeholder needs, iteratively develop data products, and manage them over time.*
- > Our survey of implementations shows there's no one way to approach data mesh. Each initiative is driven by the organization's needs and cultural readiness for the changes that it imposes.*

Recommendations

This report recommends 10 important practices to improve data mesh outcomes. Since most organizations are at the beginning of their data mesh journey, these recommendations are works in progress. However, based on implementers' common experience and collective wisdom, they can help guide your organization through the murky gap between data mesh theory and practice.

1. Evaluate how data mesh fits your organization
2. Assess business drivers
3. Map your organization's domains
4. Start small but plan for success
5. Assemble domain teams
6. Start a federated governance team
7. Define data product standards
8. Build the self-serve data platform
9. Assign data product owners
10. Employ some data mesh principles even if not adopting data mesh

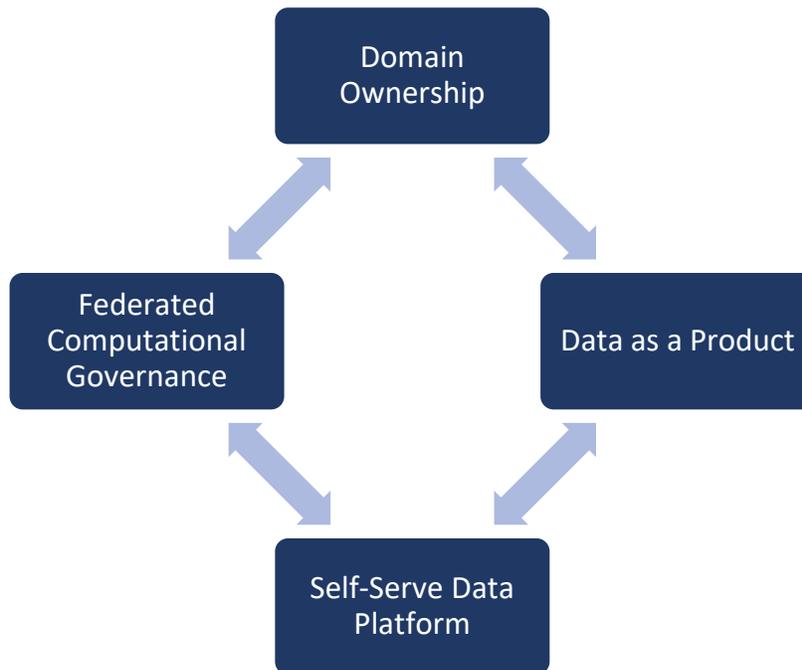
Understanding Data Mesh

Data mesh, conceived by Zhamak Dehghani, offers a new paradigm for fulfilling the promised value of data. It eschews long-standing centralized data architectures such as the data lake and the data warehouse and their associated centralized teams. These incumbent approaches weren't designed for the vast and complex data landscape that organizations manage today. In spite of the innovations that the data warehouse and data lake introduced, data remains hard to find, understand, access, and trust. Therefore, we still can't realize its full value to drive timely and accurate business decisions.

The Four Pillars of Data Mesh

To address these shortcomings, data mesh advocates for decentralizing data and distributing responsibility for it. Its approach centers on four pillars that encompass the organizational and technical changes it espouses (see figure 1).

Figure 1. The Four Pillars of Data Mesh



> **Domain ownership.** Data mesh transfers ownership of data from centralized data teams to distributed domain-oriented teams. A domain team is close to key business processes, knows the data the domain produces, and the analytics that its stakeholders need to solve problems and capitalize on opportunities. In data mesh, they are responsible for storing, cleaning, cataloging, and making their data available to its consumers in the forms they need.

- > **Data as a product.** Most organizations bring product thinking to their application development efforts. They have well-established ways of learning what operational teams need, validating designs, and developing functionality that they improve iteratively. Data mesh applies this product mindset to data in the form of a data product. Data products consist of more than just data. They include code to collect and transform data and enable managed access through APIs. They include metadata that describes the product such as schema, semantics, and quality metrics. And infrastructure to execute the product's code and store its data and metadata. Data products are scoped by their domain and delivered by a domain team.
- > **Self-serve data platform.** Domain teams need a self-serve platform to deliver and manage data products. They need to provision storage and compute infrastructure, build, deploy, and manage data product versions, clean and transform data, provide secure data access, and comply with policies and regulations. And they need to do all these things autonomously to prevent new bottlenecks from forming. Data consumers need the self-serve platform to find and access data products. To confidently use a data product, they need to understand its purpose, the timeliness of the data, and its trustworthiness. These attributes are part of the data product itself, but the platform needs to surface them for consumers to evaluate how well a product meets their needs.
- > **Federated computational governance.** Data mesh is all about distributed data and decentralized responsibility. However, there are enterprise-wide standards that cut across domains. Therefore, data mesh calls for a federated body consisting of representatives from domain teams and those with global data responsibilities such as regulatory compliance and managing quality. They work together on common concerns such as what constitutes quality, data classifications and how to handle different levels of sensitivity, modeling data that span domains, and standards for data product metadata. These are familiar data governance concerns. And most data governance programs have a degree of federated representation. However, with data mesh, it's the domain teams that apply governance standards and policies by computational, i.e., automated controls and observability into their data products.

What is a Data Product

In general terms, a data product is a reusable data asset designed for a particular use and delivered according to agreed-upon standards and schedules. An interactive dashboard, a SQL view, an embedded recommendation engine, or a fraud detection model are all examples of data products.

But a data product for data mesh is different because it calls for a particular implementation approach, at least according to the formal definition created by Zhamak Dehghani. For data mesh purists, a data product is a self-contained object that consists primarily of code that collects and transforms data, defines its metadata, and its infrastructure. The data product object must make itself discoverable, addressable, trustworthy, self-describing, interoperable, and secure primarily through APIs. The

encapsulation of its data, metadata, and infrastructure allows it to exist as a node in the mesh independent of any other data product. It's a complete package.

Also, you need more advanced technical skills to build a data mesh data product. While a dashboard is often created by a data analyst adept at wrangling and visualizing data with low-code/no-code tools, it takes a data engineer to develop a data mesh data product.

Data as a Product vs Data Products

The term data product has been around for many years. Data mesh reintroduces this idea with a much more specific definition. Beneath the notion of a data product is the foundational concept of data as a product. The former is a deliverable, the latter is a mindset. It's important to adopt this mindset before creating the deliverable.

A data product is a deliverable; data as a product is a mindset.

Paul Rankin, Head of Data Management Platforms at Roche Diagnostics, explains how they started their data mesh journey with a mindset change, "We asked ourselves, what would we need to change to move from a project-oriented, data mart-and-dashboard type of organization to a reusable data product type of organization. We knew that data as a product is really the key to delivery."

In many organizations, domain teams already practice product thinking to develop the functional applications that support their domains' business. These teams must apply product thinking to their data as well. They use product management methods in application development, such as learning what stakeholders need, refining features with constant feedback from users, and managing the product life cycle. Domain teams must also employ these approaches to their data products.

Shift Left

Data mesh paints a very different picture of how to manage data for analytics. "Shift left" is a metaphor that describes the key difference in terms of a data flow that we read from left to right. It means moving data management functions to the left where data is created and where teams with business domain knowledge and familiarity with the operational systems work. By contrast, most organizations traditionally concentrate data management functions for analytics at the righthand end of the data flow in centralized solutions such as data lakes or data warehouses. Data mesh makes them part of a domain's data products.

Data mesh shifts data management to the left toward teams with business domain knowledge and familiarity with the domain's operational systems.

Technology Changes

Figure 2 illustrates how shifting left enables self-contained data products with data, services, and infrastructure that previously existed as part a centralized analytics data management platform.

Figure 2. Shift Left from Centralized Data to Self-Contained Data Products



While shifting left is a foundational concept in data mesh, most organizations don't enable every domain to fully define its own technology stack as pictured above. As Paul Rankin explains, Roche maintained central control over the design and implementation of its self-serve data platform.

“For us it just doesn't make sense that every domain goes out and sources its own vendors, tools, and technologies,” says Rankin. “So, we made a conscious choice to make it more of a prescriptive tech stack with us building the integrations. It was still self-service, but it was more configuration type self-service that we said, okay, this is how you register your product, this is how you serve your SLOs, your SLAs, this is how you orchestrate a release, et cetera.”

Organizational Changes

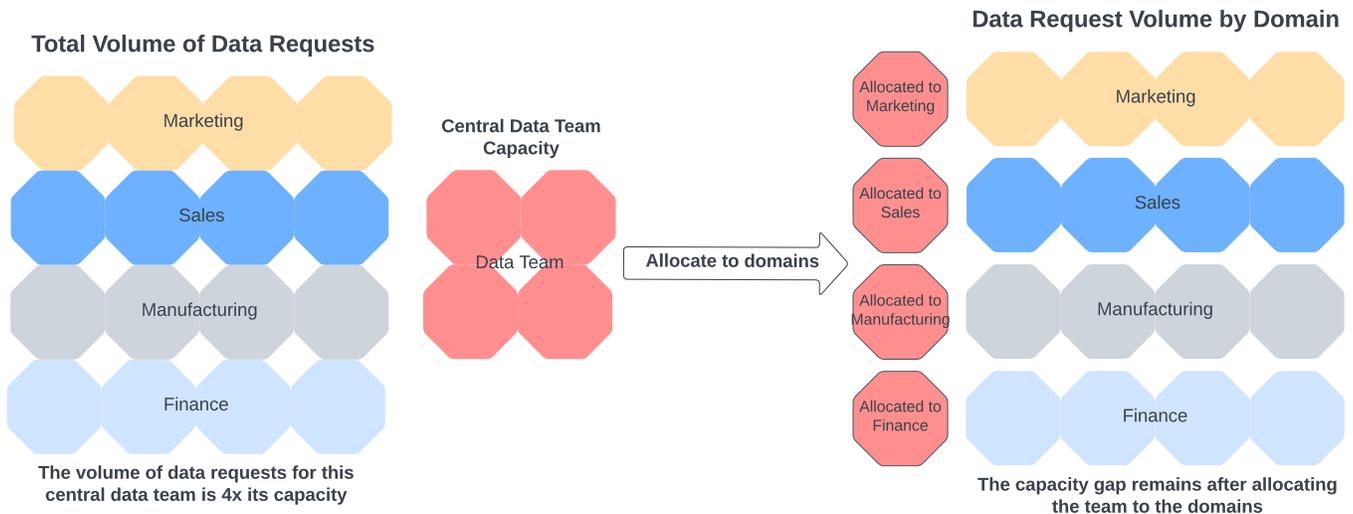
Shifting these functions to the left is not just a technology challenge. It's also organizational change that represents a fundamental departure from the status quo. As Zhamak Dehghani says in her book, *Data Mesh: Delivering Data-Driven Value at Scale*,

“Operational teams still perceive their data as a byproduct of running the business, leaving it to someone else, e.g., the data analytics team, to pick it up and recycle it into products.”

Domain ownership and data as a product require a change in mindset. Engineering teams must expand their focus beyond providing their customers with the best possible application functionality. To be domain data owners, they must add to their already-full plates the needs of data consumers.

Many companies struggle with how to shore up domain teams with additional people and skills. Reassigning members of the central data team is an option. However, in most organizations, demand for that team’s time far exceeds its capacity. Allocating them to the domains may not help because backlogged data requests must also be transferred to the domains that made them (see figure 3). Therefore, organizations must plan for supporting domain teams in their expanded role as data owners.

Figure 3. Centralized or Distributed, the Data Skills Capacity Gap Remains



Who is Data Mesh For?

Many organizations are well on their way to decentralizing data at the technical level. Modern data environments with diverse formats and distributed storage and compute tend to force organizations down the decentralized path. But the principles of domain ownership and data as a product call for people to think and work differently. Driving changes in behavior around data is one of data mesh’s main strategic objectives.

Data mesh is not for every organization. You need sufficient data size and complexity to justify the investment in a data mesh program.

Shifting left is hard work. It’s a journey, not an event. Organizations with more of the following characteristics will find the journey less fraught with peril.

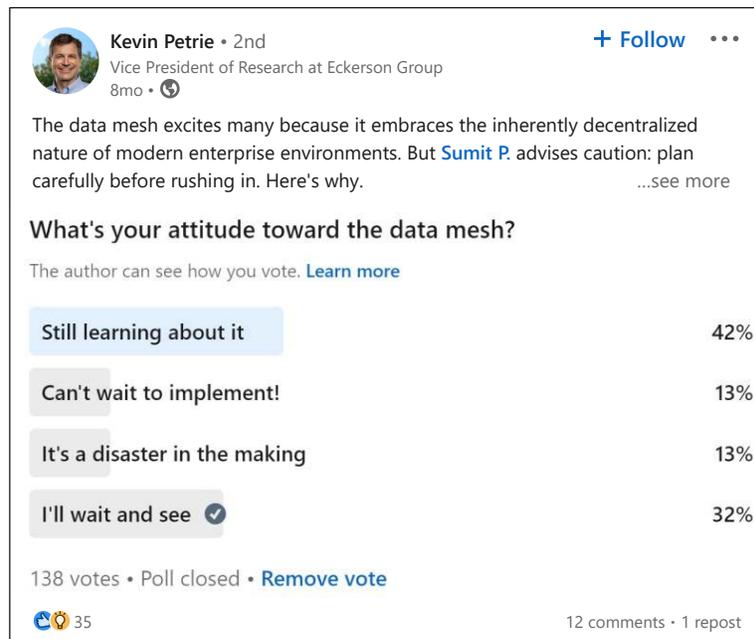
- > **Big enough and complex enough.** Data mesh is overkill for smaller organizations. To justify the effort and cost of transitioning to data mesh, an organization's data must have sufficient mass and complexity. For example, a national restaurant chain with hundreds of locations, multiple customer constituencies, marketing strategies, suppliers, talent pools, and operating schedules is big and complex enough to potentially realize benefits from data mesh. An individual owner-operated restaurant is not.
- > **Decentralized decision-making.** Organizations that tend to empower employees with decision-making authority within a framework of accountability will experience less friction in implementing data mesh. By contrast, organizations that make decisions through a multilevel approval process may find data mesh principles harder to embrace because they run counter to their existing culture.
- > **Time and money.** As we've said, data mesh is a journey. It takes time to make the fundamental changes that data mesh calls for. It also takes money to shore up the technical talent and tools needed to build data products, a self-serve data platform, and the global guardrails of federated governance.
- > **Active executive support.** Data mesh, like any change initiative, requires active executive support to be successful. This means clear, consistent messaging about what the end state looks like, the challenges to overcome, and the opportunities to seize. It also means material support through allocation of resources and a commitment to the long haul.
- > **Strong data culture.** Many organizations are committed to making decisions and driving actions with data but are hampered by legacy technology and processes. These organizations are better prepared to transition to data mesh. They understand what they can achieve if they remove the barriers that limit their ability to derive insights from data. They also recognize how effective data governance is critical for trustworthy data and have mature processes in place to define and manage data quality. On the other hand, organizations that don't rely heavily on data-driven decision-making and don't master data governance are less likely to see a worthwhile end state that they'd commit to working toward.
- > **Collaborative.** Developing data products is a collaborative process between domain teams and data consumers. Organizations whose cultures espouse and reward cross-functional collaboration are therefore in a better position to adopt data mesh. Organizations that struggle with collaboration operate in functional silos that also tend to create data silos. This results in two interrelated barriers for a data mesh initiative to overcome—breaking down functional silos and untangling their related technical debt.
- > **Ample technical skills.** Designing and building data products is a technical endeavor requiring software engineers, data engineers, infrastructure engineers, and data governance experts. Companies with a large technical staff have more to work with in assembling domain data teams. This gives them flexibility to recruit team members with the right skills and a spirit of bold curiosity for the adventure ahead.

> **Agile practitioner.** Many of the principles and practices of data mesh align with the agile development methodology. Data mesh and agile both rely on empowered product teams, collaboration, iterative development, and strong product owners. So organizations that practice agile development already have experience with the skills and methods they'll need to deliver a data mesh.

Best Practices

It may be a bit early to call what follows best practices. Many organizations are waiting on the sidelines to see how data mesh plays out. Eckerson Group ran a poll on LinkedIn, the results of which illustrate this fact. The majority of respondents, 74%, are learning or waiting and watching (see figure 4).

Figure 4. Eckerson Group LinkedIn Followers' Attitudes Toward Data Mesh



So if data mesh is a journey, at this point most of us have only walked down the block to the corner. Among those who have launched a data mesh initiative, we see common experiences and patterns based on proven success factors for any technology and organizational change program. Here are ten best practices that have already emerged from early adopters of the data mesh.

1. Evaluate How Data Mesh Fits

In the section above, **Who is Data Mesh For**, we provide criteria by which to evaluate how well data mesh fits your organization's culture, data challenges, and data maturity. But not all the criteria are of equal importance (see table 1). Let's look at the top three.

Table 1. Data Mesh Fit Criteria and Importance

Evaluation Criteria	Importance
Big, complex data	Critical
Decentralized decision-making	Very Important
Time and money	Very Important
Active executive support	Important
Strong data culture	Important
Collaborative culture	Important
Ample technical skills	Important
Agile practitioner	Important

Most important is the size and complexity of your data. If you’re dealing with the volume, velocity, and variability of big data, then consider data mesh. However, you should not pursue data mesh if you only have a handful of sources, one or two data formats, and low transaction volumes. If you’ve got the data challenges that data mesh targets, then time and money is your next concern. Doing data mesh requires material support and a commitment to the long haul. Start small with a limited pilot that shows the value data mesh will deliver. If that wins approval for further support, expand the program incrementally.

While time and money are very important, so is your organization’s culture regarding decision-making. To be nimble and responsive, domain teams must be empowered to chart their own course, change priorities, and adopt new development methods. If that goes against the grain of your organization’s culture, consider how you can overcome resistance or whether the cultural immune system will see a foreign body that it must eliminate.

2. Assess Business Drivers

Most companies struggle to leverage their data to identify opportunities, drive business decisions, and improve efficiency. These are all good reasons to consider data mesh. However, the more specifically you understand business drivers in terms of opportunities and pain points, the better you can focus your efforts on the most impactful issues. In other words, dig into what problems you’re trying to solve, how they relate to corporate strategies and objectives, and what outcomes you expect when you solve them. This provides a basis for determining if investing in data mesh is worth it. And it allows you to assess the value of your data mesh program in terms of business outcomes.

The more specifically you understand business drivers in terms of opportunities and pain points, the better you can focus your data mesh efforts.

Jules Marshall, Director of Product Data at the BBC leads their data mesh initiative. She started with a focus on business outcomes by asking some basic questions, “What do we need to get from our data? What value do we derive from it? How do we structure ourselves for long term success? That’s what we thought about when we first started considering data mesh.”

3. Map Your Organization’s Domains

It’s important to understand the domain structure of your organization at the beginning of the data mesh journey. Domains can be functional or cross-functional areas of the business. Domains should follow the existing contours of the business rather than introduce a different structure. This provides an overview of the nodes in your data mesh and a roadmap for building it out.

Paul Rankin says Roche started with the organizational structure the company was familiar with. “You have your business domains, and they’re pretty clear,” says Rankin. “Manufacturing, commercial, supply chain, finance, these are the business domains. They’re probably similar to any other type of pharmaceutical or life sciences company.”

Divide or combine domains. Companies map their domains in different ways. Some start with larger domains and break them down later as they learn details about data requirements. Other companies combine two or three functional areas into one data domain for efficiency in serving overlapping analytical needs. “The most common solution I’ve seen is this type of data is handled within the centralized data platform group or as part of an enterprise data engineering team” says Shane Murray, Field CTO at Monte Carlo. “And this is the model I used at the Times, where we decided it doesn’t make sense for the data platform group to divest ourselves of this data because they have the best visibility across the organization, and the data products are fundamentally cross-domain in their design and use.”

Handling master data. What about data subject areas that span multiple domains, such as customer or product information, a.k.a. master data? Many companies have teams that specialize in universal subject areas. “The most common solution I’ve seen is this type of data is handled within the centralized data platform group or as part of an enterprise data engineering team” says Shane Murray. “And this is the model I used at the Times, where we decided it doesn’t make sense for the data platform group to divest ourselves of this data because they have the best visibility across the organization, and the data products are fundamentally cross-domain in their design and use.”

Domain boundaries will change. However you initially map domains, plan for the boundaries to change as experience reveals better ways to organize them. Domain changes are driven by other forces as well such as changes in business strategy, new products, re-orgs, mergers and acquisitions, divestitures, and economic conditions.

Plan for domain boundaries to change as experience reveals better ways to organize them.

So don't spend a lot of time striving for perfection. Consider Paul Rankin's observations about defining domains. "What we found over the last year or two, is that we could spend hours arguing about what a domain is and who owns it and who's responsible for it. But does it really matter at the end of the day to the value that a data product provides? Probably not. What matters is clear business ownership, that is key, and a durable team to build and manage products."

4. Start Small but Plan for Success

"Don't try to boil the ocean." "Don't do a big bang release." These pearls of wisdom apply to just about every technology initiative. Data mesh is no different. "What does that transition plan look like to get there longer term? That's how I approach this," says Jules Marshall. "This isn't a big bang like, we can have decentralized data across the organization all at once. Being use case driven is essential to ensure we're delivering business value as we go, continually bringing more of the organization into the new ecosystem of data products." Start small with one or two domains and data use cases that will deliver clear value.

Start small with one or two data use cases that will deliver clear value. Focus on a specific problem with a manageable scope and iterate. This enables you to pivot quickly in response to stakeholder feedback from hands-on experimentation. Thus, team members can more quickly find the maximum value a data product can deliver, or determine that a data product is not worth developing.

Shane Murray offers another observation from his work with customers. "I think another question is do we start with a data product that exists and delivers value to the business or are we seeking out some sort of new territory that is unencumbered by the past? I've seen more success when it's establishing something new. It can be tricky because then you've got to actually drive adoption. But it can be a better approach because you're not hindered by working with something that was maybe built on old tool sets, or on a prior generation of data platform infrastructure."

5. Assemble Domain Teams

Domain teams need people with business domain knowledge that understand how stakeholders can derive value from data. They should understand the domain's business processes, challenges, and opportunities. The more familiarity the team has with the domain's objectives and operations, the better equipped they are to develop impactful data solutions.

The team needs members with data engineering and analysis skills who are familiar with the domain's application and data stacks. Expertise in machine learning and artificial intelligence may also be required

depending on the domain's objectives. The team also needs members with an understanding of information security and data privacy policies and regulations to avoid compliance issues.

Domain teams also have to manage the product delivery process. They must plan their work, prioritize features, and execute, clearing roadblocks to deliver product iterations. And they must keep stakeholders involved and informed of progress. Therefore, they need members that know the delivery methodology that the team opts for, such as scrum or lean kanban.

As you assemble this cross-functional team, look for team members that are either already treating data as a product or are receptive to changes in responsibility that domain ownership entails. According to Shane Murray, "Where you have teams that are already outfitted and operating as data owners, you can start showing the value earlier."

6. Start a Federated Governance Team

A federated governance team includes representatives from each domain team and members with global data responsibilities such as regulatory compliance and managing quality. Jules Marshall believes this is critical for success "For me, it ends up coming down to two main things, strong data management and governance. Getting those practices in place, getting the governance in place. Then we can draw a line in the sand and make improvements from that point forward."

If you start small—and we hope you do—the first iterations of your federated governance team will be similar to your initial cross-functional domain teams. As you bring new domains into your data mesh program, you'll add their representatives to the federated governance body. Roche calls this group the cross-domain governance board. In addition to domain representatives, the group includes members from their IT quality and regulatory management functions. The board is responsible for defining overarching standards for issues such as data classification, appropriate data use, and regulatory compliance.

The federated governance team formulates standards and policies that apply to all data products. But the domain teams must apply those policies to each of their data products, which is not feasible to do manually. So, automation is key. For example, domain teams must apply global standards for self-documentation to their products. Some organizations are using data catalogs to harvest metadata from data products and check that required attributes are supplied. Many use data observability functions to collect data quality attributes that indicate a data product's current trustworthiness.

7. Define Data Product Standards

One of the first things federated governance should tackle is defining global standards for a minimally viable data product (MVDP). This is a set of standards describing requirements that all data products

must meet. Minimum viability defines what's good enough for stakeholders to start using a data product even if it doesn't yet have all the requested features. An MVDP is ready for iterations of development incorporating feedback from users in a cycle of continuous improvement. Plan for MVDP standards to change as you learn from experience and onboard new domains.

One of the first things federated governance should tackle is defining global standards for a minimally viable data product (MVDP).

Here are three key areas to address in defining MVDP characteristics.

- > **A defined scope.** A single data product can't do everything for everybody. Therefore, it should have a well-defined scope that addresses a specific business need or use case. Domain teams should focus first on developing data products for their domain, whose business processes and source systems they know.
- > **Well documented.** A data product must communicate key facts about itself, such as the value it delivers, how it should be used, the product owner, the lineage of its data, service level objectives, its quality measures, sample datasets, and how to address it through an API or query interface. Each organization has to determine the minimum degree of documentation that applies to all data products.
- > **Designed for continuous improvement.** Whatever your data product looks like at the beginning of its life will change. So, it should be designed for continuous improvement based on user feedback and observability data.

8. Build the Self-serve Data Platform

It might be tempting at first to just build data products. The limited scope of a data product and domain ownership can put trustworthy data in the hands of stakeholders faster. If you define and follow MVDP standards, more comprehensive metadata and documentation make it easier for data consumers to find and use data products.

However, these outcomes are unlikely to scale without a self-serve data platform. So start building the platform immediately. It must serve two constituencies, data product developers and data consumers. Developers need automation that helps them streamline the process of building, deploying, and testing new data product versions. Data consumers need to discover, understand, access, and trust data products that meet their needs.

Automation for developers. Paul Rankin describes one of his organization's early data products and what they learned about the limitations of building it manually. "At that time, we had just signed a deal

with Snowflake. We knew that we had a platform or a base that we could use. We were doing everything by hand. Then we realized that this might work with one use case. But to really scale this across Roche, we needed to automate with dataops.”

DataOps functionality is playing a key role in many data mesh implementations because it helps teams deliver product iterations faster. DataOps automates deploying code and infrastructure, a.k.a. continuous integration and deployment (CI/CD). It also automates testing, orchestrating pipelines, logging results, and alerting about issues. Data observability has also emerged as critical. It provides information about data quality—its validity, completeness, consistency, and timeliness—operational performance of processes and the infrastructure that runs them.

DataOps is playing a key role in many data mesh implementations through CI/CD and testing automation that helps teams deliver value faster.

Discovery and access for data consumers. Data products deliver value when consumers can find them and get access to them. They need to know the product’s purpose, the data’s provenance, service level objectives, quality measures, and other important attributes. Some data mesh practitioners register products in an existing data catalog, which makes them discoverable and provides metadata and documentation. But most catalogs don’t currently enable access to the data they document.

A data marketplace enables users to find data products and their metadata and documentation but also delivers the data itself. It provides approval workflows for users to request access from data product owners. Data mesh implementers haven’t embraced marketplaces yet. But this is likely to change since delivering data and streamlining access approval closes the loop that puts data in users’ hands so they can derive insights. In his blog, [Data Mesh’s Missing Ingredient: A Data Marketplace](#), Wayne Eckerson says,

“A data marketplace closes the last mile between data providers and data consumers. It represents a fourth-generation data architecture that perfectly complements a data mesh.”

Establishing the self-serve data platform capabilities early in your data mesh journey is important. However, it’s also important to not let perfect be the enemy of good. Building a self-serve data platform is a development effort in itself with its own degree of minimal viability. Design for change, reusability, and scalability. And then iterate to continuously improve.

9. Assign Data Product Owners

The role of the data product owner is not yet well established in practice. However, it's clear that it involves a blend of strategic management at the domain level and tactical management at the product level. "What we're looking for is people with that cross hash of skills, somebody with a data background as well as a product background. There's a synergy there", says Jules Marshall, "The data maturity of the organization matters. The less mature, the more you need these skills at the start."

Candidates for the product owner role should exhibit the following key characteristics.

- > **Domain experience.** As a member of the domain data team, the data product owner must have domain knowledge to effectively collaborate with stakeholders.
- > **Value driven.** A data product is worth developing and maintaining only if it delivers value. Therefore, the data product owner should be laser-focused on and highly motivated by fulfilling their data products' value proposition.
- > **Collaborative.** As the primary point of contact for a cross-functional domain team, the data product owner must love bringing people together and helping them work effectively together.
- > **Communication skills.** The data product owner works with both stakeholders and the domain team. So they must communicate clearly and with empathy, moderating their style depending on the audience.
- > **Technical knowledge.** The data product owner must be familiar with the domain's applications and data to collaborate effectively with their team's technical resources such as data engineers. They must understand the strengths and limitations of the infrastructure and tools they use to create and manage data products.
- > **Product delivery skills.** Delivering data products, whether an MVDP or subsequent iterations, requires keeping the work on track and stakeholders informed. The data product owner uses the product delivery methods of the domain, whether that's an Agile approach such as Scrum or traditional waterfall project management, to get products out the door.

The responsibilities of the data product owner put these characteristics to work in defining and executing a data strategy for their domain. The following responsibilities are critical functions of the role.

- > **Maintain the domain's data roadmap.** Develops and maintains a roadmap in collaboration with stakeholders that defines data objectives for the domain and provides guidelines for data product development priorities.

- > **Manage the backlog.** Evaluates data requests from stakeholders, clarifies requirements, and maintains a backlog of pending work ranked according to roadmap objects, value delivery, and risk mitigation.
- > **Represent stakeholders.** Draws on domain knowledge to represent stakeholders in data product ideation, design, and iterative development work. Ensures alignment with roadmap strategy and priorities.
- > **Deliver data products.** Monitors development progress, helps clear delivery roadblocks, organizes stakeholder demos and collaborative design sessions, and keeps stakeholders informed of progress.

10. Employ Some Data Mesh Principles Even if not Adopting Data Mesh

You may come to the conclusion that data mesh is not a good fit for your organization. While it's not the right approach for everyone, some data mesh concepts are worth considering under any circumstances. For example, treating data as a product keeps stakeholder needs and value delivery front and center. In data mesh, an autonomous domain team creates and owns a data product. However, even a small, centralized data team can better serve their stakeholders by applying product thinking to their deliverables.

While data mesh is not the right approach for everyone, some of its concepts are worth considering under any circumstances.

A self-serve data platform removes overhead for data product developers to accelerate their continuous build-test-deploy cycles. The platform should also make data easy for consumers to find, understand, and use, which is critical whether you adopt data mesh or not. Incorporating the dataops and data marketplace functionality that make up a self-serve data platform is a good investment.

Automated data management functions are critical capabilities whether you adopt data mesh or not. For example, you can improve data governance coverage, accuracy, and speed with an automated process that identifies and tags regulated sensitive data. Also, you should automate data access control applied across your data environment to mitigate risks of breaches and compliance issues.

Even if your organization does not adopt data mesh, investing in these universal capabilities makes sense based on the benefits they yield regardless of data architecture and methodology. Also keep in mind that your organization may evolve and become better suited to data mesh down the road. If you've developed these capabilities, you're in a much better position to implement data mesh having already done some of the work.

Conclusion

Companies have always been challenged to derive value from data. Over the years, many solutions have emerged, such as the data warehouse and the data lake, in response to evolving data technologies and business needs. We've pinned our hopes on them, often expecting them to finally and forever deliver data's full value to drive business decisions and actions. While none of them have been the proverbial silver bullet, each approach leaves us with enduring data management principles that propel the state of the art forward.

Data mesh is a current influential contributor to the art of data. It's designed for the highly distributed and diverse world of modern data environments. Its four pillars, domain ownership, data as a product, the self-serve data platform, and federated computational governance, are helping us rethink our relationship with data. It asserts that we should decentralize both the ownership of data and the data itself, shifting them left from centralized data teams toward the functional domains that create and use data to run their business.

Those implementing data mesh must work within the bounds of what their organization's culture, existing technologies, and resources will allow.

However, as we've seen in this report, practitioners implementing data mesh must work within the bounds of what their organization's culture, existing technologies, and resources will allow. Every data mesh implementation is different. If you proceed with data mesh, be prepared for a journey that unfolds iteratively according to your company's needs and readiness for the changes that data mesh imposes.

About Eckerson Group



Wayne Eckerson, a globally-known author, speaker, and consultant, formed **Eckerson Group** to help organizations get more value from data and analytics. His goal is to provide organizations with expert guidance during every step of their data and analytics journey.

Eckerson Group helps organizations in three ways:

- > **Our thought leaders** publish practical, compelling content that keeps data analytics leaders abreast of the latest trends, techniques, and tools in the field.
- > **Our consultants** listen carefully, think deeply, and craft tailored solutions that translate business requirements into compelling strategies and solutions.
- > **Our advisors** provide competitive intelligence and market positioning guidance to software vendors to improve their go-to-market strategies.

Eckerson Group is a global research, consulting, and advisory firm that focuses solely on data and analytics. Our experts specialize in data governance, self-service analytics, data architecture, data science, data management, and business intelligence.

Our clients say we are hard-working, insightful, and humble. It all stems from our love of data and our desire to help organizations turn insights into action. We are a family of continuous learners, interpreting the world of data and analytics for you.

Get more value from your data. Put an expert on your side. **Learn what Eckerson Group can do for you!**

