Literature Review on Data Mesh

Seminar paper

Bukaieva, Mariia, FH Wedel, Wedel, Germany, stud106892@fh-wedel.de

Abstract

Data Mesh has gained increasing attention as a new enterprise data management concept focusing on a decentralized approach. It is perceived as a response to the limitations of centralized data architectures, especially when organizations face challenges caused by growing data volumes, diverse data sources, and many business-specific data use cases. This seminar paper reviews the current academic and industry-oriented literature on Data Mesh and examines which key dimensions characterize the state of research. Sixteen selected publications were analyzed using a concept-centric method. The review identifies five recurring dimensions: architectural components, organizational capabilities and roles, governance mechanisms, technical capabilities, and implementation challenges. The findings show that successful implementation depends not only on technical architecture, but also on organizational maturity, clear responsibilities, governance structures, platform support, and suitable incentives. Data Mesh appears to be a promising but still developing concept. More empirical research is needed to better understand its long-term outcomes, practical limitations, and suitability for different types of organizations.

Keywords: Data Mesh, Data Governance, Data Products, Decentralized Data Architecture.

Table of Contents

1 Introduction

Nowadays, data has become a crucial resource for organizations. It supports decision-making, analytics, and data-driven innovation (Azeroual & Nacheva 2023; Blohm et al. 2024). In order to support massive data loads, organizations have used centralized approaches like data warehouses, data lakes, and data lake houses to store, process, and analyze data (Bode et al. 2024; Gieß & Hutterer 2025; Palanisamy 2025). However, these approaches face various limitations when data volume, data variety, and the number of data use cases grow (Bode et al. 2024; Gieß & Hutterer 2025).

Another important limitation that can often become a major bottleneck are central data teams. They are often responsible for collecting, preparing, and providing data for many different business domains (Bode et al. 2024). At the same time, these teams may lack the detailed domain knowledge needed to understand the meaning and quality of specific data sources (Bode et al. 2024). Therefore, the problem is not only technical. It is also connected to ownership, responsibilities, and organizational structure. This also matters for managers because such bottlenecks slow decision-making and data-driven innovation across the whole organization.

Data Mesh concept, introduced by Dehghani (2019), has been proposed as a response to all these challenges. It is a decentralized and socio-technical approach to data management that organizes data around business domains and gives domain teams more responsibility for their own data (Dehghani 2019). The popularity of this concept can be seen in Figure 1, where it shows that since 2022 this term has gained more and more attention on the internet.

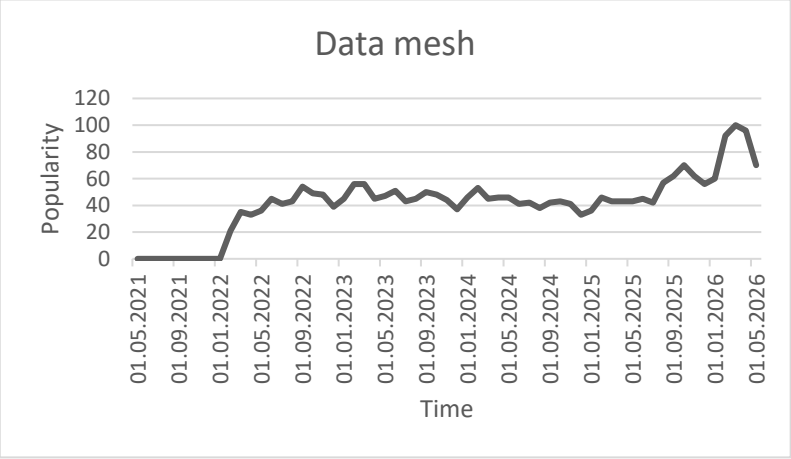

Figure 1. Trending of the topic "Data Mesh" (Google Trends, visited on 24/05/2026).

Although Data Mesh is discussed as a promising approach, the literature also shows that it is still a young concept. Winter and Hackl (2023) state that Data Mesh still lacks a comprehensive scientific background. Bode et al. (2024) also point out that empirical insights into Data Mesh implementations are still limited. This makes it relevant to examine how current academic and industrial literature discusses the concept.

Therefore, this paper addresses the following research question:

What is the state of the art of academic and industrial literature on Data Mesh and which key dimensions characterize current research?

To answer this question, I review selected academic and industry-oriented publications on Data Mesh. The next chapter provides the background needed to understand the concept. After that, I describe the literature review process. Then, I analyze the selected literature according to the key dimensions identified during the review. Finally, I discuss commonalities, differences, practical implications, and research gaps in the selected literature.

2 Background

2.1 Data Mesh as a decentralized data management approach

Traditional enterprise data management is often based on centralized architecture (Whyte et al. 2022). In this architecture, data from different operational systems is collected and managed by central analytical or IT teams (Bode et al. 2024; Dolhopolov et al. 2024). Data Mesh changes this logic by distributing data ownership to different business domains (Panigrahy et al. 2023; Gieß & Hutterer 2025).

A domain can be understood as a business area such as sales, marketing, finance, or customer support. In Data Mesh architecture, such domains are responsible for managing and sharing their own data (Dehghani 2019). The main idea is that the teams closest to the data are also best positioned to understand its meaning, quality, and use it (Bode et al. 2024; Palanisamy 2025).

2.2 Core principles of Data Mesh

Data Mesh is commonly described through four core principles: domain-oriented data ownership, data as a product, self-serve data infrastructure, and federated governance (Dehghani 2019).

First, domain-oriented data ownership means that business domains take responsibility for their own data. This shifts ownership away from a single central data team and toward the teams that produce and understand the data (Panigrahy et al. 2023; Gieß & Hutterer 2025). In Figure 2 it is visible that in a federated data governance context, there is no strict hierarchy among different domains. This allows each team to follow their local policy.

Second, data as a product means that data should be prepared and maintained for consumers. Data is not treated only as a technical by-product, but as something that should be usable, trustworthy, and reusable (Blohm et al. 2024; Dolhopolov et al. 2024).

Third, self-serve data infrastructure means that domain teams should receive platform support for creating, managing, and sharing data products (Dolhopolov et al. 2024). This avoids that every domain has to build the same technical capabilities on its own. Self-service data platform is the core of the federated data governance as seen from Figure 2.

Fourth, federated governance means that domains keep autonomy, while the organization still defines common rules and standards (Dolhopolov et al. 2024; Winter & Hackl 2023). This principle is important because Data Mesh needs both decentralization and coordination.

Together, these principles provide the conceptual basis for the literature review.

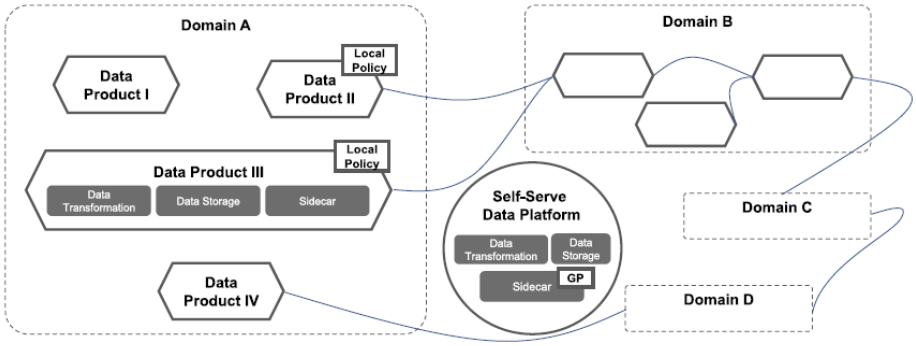

Figure 2. *Federated Data Governance (Bode et al. 2024).*

Note. This figure shows different levels of granularity (high on the left and low on the right).

3 Literature Review Methodology

In this chapter, I describe how I identified, selected, and analyzed the literature for this review. I followed a structured literature review approach based on the grounded theory literature review method proposed by Wolfswinkel et al. (2013). The method divides the review process into the stages define, search, select, analyze, and present.

3.1 Define

In the first step, I defined the scope of the literature review and the criteria for selecting relevant sources. The focus of this paper is Data Mesh, and which key dimensions characterize current research about it. Therefore, I included publications discussing Data Mesh principles, governance, data products, decentralized ownership, implementation strategies, and organizational implications.

I only included English-language publications published between 2021 and 2026. I used this time frame because Data Mesh is still a relatively new concept and most relevant publications were published in recent years. I mainly focused on peer-reviewed journal articles and conference papers. However, I also included selected industry publications because the topic is strongly influenced by practitioners and real-world implementations.

3.2 Search

In the exploratory phase I used ChatGPT Deep Research as a supporting tool to identify possible directions and search terms. The final selection of literature was based on manual searches in scientific databases and digital libraries. The main databases I used were Google Scholar, IEEE Xplore, SpringerLink, ScienceDirect and Wiley. I also used ResearchGate and arXiv to identify newer publications and ongoing discussions related to Data Mesh.

To organize and manage the literature, I used Mendeley throughout the review process.

I started with a broader exploratory search phase. During this phase, I experimented with different keywords to better understand the research field and identify recurring concepts. The main search terms I used were “data mesh”, “data management”, “decentralized data architecture”, “data as a product”. Later, I refined the search using more focused keyword combinations. These included “data mesh AND data product”, “data

mesh AND architecture", "data mesh AND implementation". I mainly searched for these terms in titles, abstracts, and keywords. Another useful way to identify suitable papers was to examine the references of already selected papers.

3.3 Select

I selected the literature in multiple stages. In the first stage, I collected around thirty potentially relevant publications during an exploratory search phase. I evaluated the identified publications based on three main criteria. First, the publication had to fit the research question and focus clearly on Data Mesh. Second, the publication needed sufficient scientific credibility. Third, the publication had to contribute meaningful insights related to architecture, governance, implementation, or organizational aspects of Data Mesh. I reduced the number to around twenty publications.

In the second stage, I manually searched for additional resources in various online libraries and journals. I was able to collect twenty-eight papers in total. I excluded several types of sources during the review process. First, I excluded eight bachelor's and master's theses because they are usually not peer-reviewed and often have lower scientific credibility. Second, I excluded two books and two non-scientific publications without sufficient references or citations. Many of these books mainly reflected the perspective of individual authors and were difficult to validate academically. Finally, I excluded a couple of publications unrelated to enterprise data management or publications using the term "mesh" in a different technical context

In the final stage, I added selected industry publications and implementation examples. These included consulting publications such as PwC articles, industry-oriented articles involving IBM employees, and implementation examples such as Kuehne+Nagel. I included these sources because academic literature on Data Mesh is still limited and many practical insights currently come from industry.

After the final filtering process, sixteen publications remained and formed the basis of this literature review.

3.4 Analyze

I analyzed the selected literature qualitatively using a concept-centric approach inspired by grounded theory literature review methods (Wolfswinkel et al. 2013). During the analysis process, I applied three coding stages: open coding, axial coding, and selective coding.

First, I used open coding by reading the papers and identifying recurring concepts, themes, and implementation patterns related to Data Mesh. Examples of recurring concepts included domain ownership, data products, federated governance, metadata management, self-service platforms, data contracts, and implementation challenges.

Second, I used axial coding to connect related concepts and group them into broader categories. For example, concepts such as decentralized accountability, domain ownership, and team responsibility were grouped into organizational ownership and governance-related categories.

Finally, I used selective coding to derive the higher-level dimensions that structure this literature review. Through this process, the final dimensions of the review emerged iteratively from the literature. These dimensions included architectural components, organizational capabilities & roles, governance mechanisms, technical capabilities and implementation challenges.

3.5 Limitations

This methodology has several limitations. First, some relevant publications could not be included because access required subscriptions, memberships, or additional payment through databases such as Springer, Wiley, ResearchGate and others.

Second, the review focused only on recent publications after 2021. As a result, older publications related to decentralized data architectures may not have been included.

Another limitation is the current maturity of the research field itself. Data Mesh is still a relatively young concept and much of the available literature is practitioner-oriented rather than empirically validated. Many publications are conceptual discussions instead of long-term empirical studies. Therefore, some findings should be interpreted carefully.

Lastly, since Data Mesh concept is mostly business driven, a lot of resources are written from consulting perspective or as business reports. They are written by people from industry, by students or independent authors. Thus, such sources were only included when they provided relevant practical insights and when their role as industry-oriented literature was made explicit.

4 Findings and Results

4.1 Overview of the Concept Matrix

In the final analysis, I examined 16 selected publications on Data Mesh. To structure the findings, I created a concept matrix following the concept-centric approach of Webster and Watson (2002). The matrix groups the reviewed literature according to five recurring dimensions: architectural components, organizational capabilities and roles, governance mechanisms, technical capabilities, and implementation challenges.

Paper	Architectural Components	Organizational Capabilities & Roles	Governance Mechanisms	Technical Capabilities	Implementation Challenges
Machado et al. (2022)	●		●	●	
Azeroual & Nacheva (2023)	●		●		●
Winter & Hackl (2023)		●	●		●
Bode et al. (2024)	●	●	●		●
Kernstock et al. (2024)	●	●	●	●	●
Sisson et al. (2024)		●	●		●
Dolhopolov et al. (2024)			●	●	
Blohm et al. (2024)		●			
Gieß & Hutterer (2025)		●	●		●

Lew et al. (2025)					●
Rajasekaran & Raj (2025)	●		●	●	●
Pakrashi et al. (2023)	●	●	●	●	●
Palanisamy (2025)	●			●	
Panigrahy et al. (2023)	●		●	●	
Baião and Marconi (2024)	●	●	●	●	●
Whyte et al. (2022)	●	●	●		●

Table 1. Concept Matrix

A marked cell indicates that the paper discusses the respective dimension as a central theme. In the following subsections, I describe each dimension and its main sub-categories.

4.2 Architectural Components

The first dimension covers the architectural components of Data Mesh. As mentioned above, most papers explain Data Mesh through its four core principles: domain-oriented ownership, data as a product, self-serve data infrastructure, and federated computational governance (Panigrahy et al. 2023; Blohm et al. 2024; Bode et al. 2024; Dolhopolov et al. 2024). Figure 3 clearly shows that these four principles are interdependent.

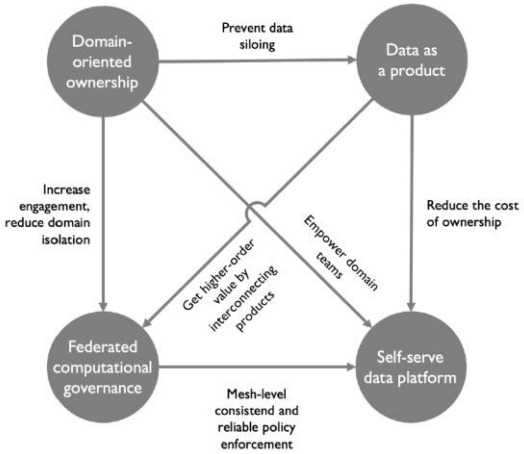

Figure 3. Four principles of Data Mesh and how one requires another (Winter & Hackl 2023).

And a central idea is to move away from one centralized data platform toward a more distributed structure. In this structure, data is organized around business domains instead of being managed mainly by one central data team (Machado et al. 2022; Azeroual & Nacheva 2023; Panigrahy et al. 2023). Domains are therefore one of the main building blocks of Data Mesh.

Another important component is the data product. In the literature, data products are described as reusable units that should be easy to find, understand, trust, access, and reuse (Bode et al. 2024; Blohm et al. 2024). The industrial literature often makes this

more concrete. For example, the Kuehne+Nagel case describes a data product as a combination of data, code, interfaces, documentation, ownership, and usage agreements (Baião & Marconi 2024). This shows that a data product is more than just a dataset.

The self-serve data platform is another recurring component. Its purpose is to support domains in creating, publishing, finding, and using data products (Winter & Hackl 2023; Blohm et al. 2024). This platform is important because decentralized ownership would be difficult if every domain had to build all technical capabilities by itself. Therefore, shared tools and infrastructure can reduce duplicated work across domains (Blohm et al. 2024; Dolhopolov et al. 2024).

Thus, the reviewed literature presents Data Mesh as an architectural shift that depends strongly on ownership and governance arrangements.

4.3 Organizational Capabilities and Roles

The second dimension covers organizational capabilities and roles. One of the most important aspects of Data Mesh is the change in data responsibility (Whyte et al. 2022). Instead of keeping most responsibilities in central data or IT teams, Data Mesh shifts ownership toward domain teams and subject-matter experts (Panigrahy et al. 2023; Winter & Hackl 2023; Bode et al. 2024).

Several papers explain why this shift is needed. Central data teams can become bottlenecks when many departments need data at the same time. They may also lack the domain knowledge needed to fully understand the data they manage (Blohm et al. 2024; Bode et al. 2024; Whyte et al. 2022). Domain teams are closer to the meaning and use of their data. Therefore, they are expected to be better able to judge data quality and provide useful data products (Azeroual & Nacheva 2023; Bode et al. 2024).

The literature also shows that Data Mesh requires a product-oriented way of working. Domain teams should not only provide data once for a single project. They should maintain data products for consumers over time (Machado et al. 2022; Panigrahy et al. 2023). Winter and Hackl (2023) describe data products as artifacts that include data, code, metadata, policies, and infrastructure dependencies. This means that ownership includes the full lifecycle of a data product.

At the same time, this shift is not easy. Domain teams may not have the needed skills, resources, or motivation to take over data product ownership (Bode et al. 2024; Lew et al. 2025). The Kuehne+Nagel case also shows an incentive problem: the team that produces data is not always the team that benefits from better data quality (Baião & Marconi 2024). This can make it difficult to motivate data producers to invest time in documentation, quality, and maintenance.

Therefore, Data Mesh requires more than new roles on paper. It also requires skills, incentives, and collaboration between domain teams, platform teams, and governance roles.

4.4 Governance Mechanisms

The third dimension covers governance mechanisms. Governance is central because Data Mesh tries to combine two goals. On the one hand, domains should have more autonomy. On the other hand, the organization still needs common standards. The literature usually describes this as federated governance or federated computational governance (Winter & Hackl 2023; Dolhopolov et al. 2024; Bode et al. 2024).

Federated governance means that domains can make local decisions, but still follow shared rules across the organization (Winter & Hackl 2023; Blohm et al. 2024). These rules can relate to data quality, security, privacy, access rights, interoperability, compliance, and ownership (Azeroual & Nacheva 2023; Kernstock et al. 2024). This is important because decentralization without standards could create new silos and inconsistencies (Azeroual & Nacheva 2023; Panigrahy et al. 2023).

Some authors focus especially on the computational part of governance. Dolhopolov et al. (2024) describe federated computational governance as a combination of collaboration on schemas, semantics, and lineage, together with tools that enforce security, transparency, and legal policies. This means that governance should not only exist as written rules. It should also be supported by technical mechanisms such as metadata systems, policy enforcement, access control, and data contracts (Dolhopolov et al. 2024).

Industry-oriented literature also confirms the importance of governance. The Kuehne+Nagel case mentions access approval, ownership, usage agreements, metadata, and quality checks as practical governance elements (Baião & Marconi 2024). Bode et al. (2024) also identify the transition toward federated data governance as one of the main challenges in Data Mesh implementation.

To conclude, the governance dimension shows that Data Mesh does not remove the need for coordination. Instead, Data Mesh changes how coordination is organized. Governance becomes the mechanism that connects domain autonomy with enterprise-wide consistency.

4.5 Technical Capabilities

The fourth dimension covers the technical capabilities needed for Data Mesh. Although Data Mesh is not only a technical concept, the literature shows that technical infrastructure is necessary to make it work in practice (Blohm et al. 2024; Dolhopolov et al. 2024; Palanisamy 2025).

One important capability is metadata management. Metadata helps users find, understand, access, and evaluate data products (Azeroual & Nacheva 2023; Blohm et al. 2024; Dolhopolov et al. 2024). Dolhopolov et al. (2024) describe metadata management as a core component of federated governance in Data Mesh platforms. They also mention requirements such as semantic enrichment, indexing, lineage, versioning, usage tracking, contract management, and policy enforcement.

Data catalogs are also often mentioned. They provide an overview of available data products and support discovery across domains (Azeroual & Nacheva 2023; Blohm et al. 2024). In practice, catalogs can be combined with automation, access management, observability, quality checks, and user management (Baião & Marconi 2024).

Another technical capability is the use of APIs and standardized interfaces. Bode et al. (2024) state that data products should include metadata and access options such as APIs. Pakrashi et al. (2023) also emphasize APIs and common protocols in their sector-specific Data Mesh for the dairy industry. This shows that technical standardization is needed so that data products can be used across domains.

In summary, technical capabilities are not the main idea of Data Mesh, but they are necessary enablers. Without metadata, catalogs, APIs, automation, and platform services, decentralized data ownership would be difficult to operate.

4.6 Implementation Challenges

The implementation dimension shows that Data Mesh adoption is not merely an architectural migration, but an organizational transformation requiring management support, new capabilities, and gradual change. Figure 4 illustrates such a transformation journey.

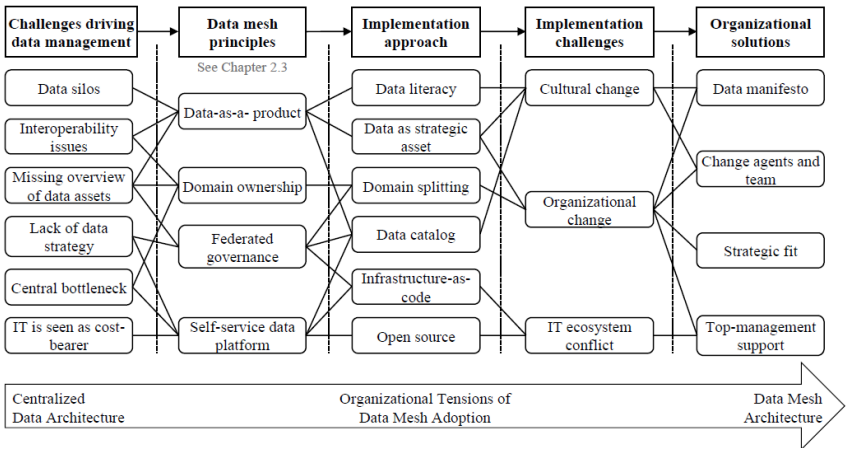

Figure 4. Alpha's Data Mesh transformation journey (Kernstock et al. 2024).

Several papers point out that Data Mesh is still a young concept and that empirical knowledge is limited (Machado et al. 2022; Winter & Hackl 2023; Bode et al. 2024; Kernstock et al. 2024). Bode et al. (2024) state that more research is needed on motivations, challenges, implementation strategies, impact, and archetypes. Kernstock et al. (2024) also highlight the need for more empirical understanding of transformation processes and tensions.

One major challenge is organizational change. Data Mesh requires domain teams to take responsibility for data products, but these teams may lack the necessary skills, motivation, or resources (Bode et al. 2024; Lew et al. 2025). This is connected to cultural resistance and the difficulty of distributed accountability (Lew et al. 2025; Rajasekaran & Raj 2025). Bode et al. (2024) also warn that organizations may adopt Data Mesh because of hype, instead of having a clear motivation.

Another challenge is coordination. Data Mesh decentralizes ownership, but domains still need to coordinate around standards, quality, security, and interoperability (Azeroual & Nacheva 2023; Dolhopolov et al. 2024). This makes federated governance difficult in practice. Organizations need to avoid both extremes: too much central control and too much fragmentation (Sisson et al. 2024).

There are also technical and resource-related barriers. Dolhopolov et al. (2024) argue that Data Mesh still lacks sufficient technology support for widespread adoption. Winter and Hackl (2023) mention high initial investments and enterprise-level trade-offs. Lew et al. (2025) also report platform engineering investment, longer implementation timelines, and mixed outcomes.

Overall, the implementation dimension shows that Data Mesh is not suitable for every organization in the same way. Its success depends on organizational maturity, domain capabilities, governance design, platform support, and clear motivation (Bode et al. 2024; Lew et al. 2025; Rajasekaran & Raj 2025).

5 Discussion

5.1 Commonalities, Differences, and Main Tensions

The reviewed literature shows a relatively stable conceptual core: most publications build on the same four Data Mesh principles. However, they differ in how strongly they emphasize architecture, governance, technical infrastructure, or organizational transformation.

Some authors mainly present Data Mesh as a new data architecture concept and compare it with data warehouses, data lakes, and data lake houses (Machado et al. 2022; Azeroual & Nacheva 2023; Panigrahy et al. 2023). Other authors focus more on governance, metadata, policy enforcement, and platform requirements (Dolhopolov et al. 2024; Pakrashi et al. 2023; Palanisamy 2025). A third group of papers is more interested in practical adoption and organizational change, especially in large organizations or concrete company cases (Winter & Hackl 2023; Bode et al. 2024; Kernstock et al. 2024; Baião & Marconi 2024; Whyte et al. 2022).

One important difference is also the scope of the papers. Gieß and Hutterer (2025) compare Data Mesh with other modern data management approaches such as data platforms, data spaces, and data fabrics. Blohm et al. (2024) also discuss Data Mesh together with data products and data fabric. Sisson et al. (2024) go even further by connecting Data Mesh to data intermediaries, data supply chains, co-operation, and co-regulation. These papers show that Data Mesh is not discussed in isolation, but as part of a new approach in data management.

The main tension in the literature is the balance between decentralization and standardization. Data Mesh gives domains more autonomy, but this autonomy can create new problems if common standards are missing. Therefore, the literature repeatedly emphasizes the need for governance, metadata, interoperability, and shared platform capabilities (Winter & Hackl 2023; Dolhopolov et al. 2024; Bode et al. 2024). This means that Data Mesh does not remove the need for coordination. It mainly changes how coordination is organized.

A second tension concerns the promise and the complexity of Data Mesh. The concept is often presented as a way to reduce bottlenecks and improve scalability. However, several papers also show that implementation can create new difficulties, especially around skills, incentives, ownership, and governance (Bode et al. 2024; Kernstock et al. 2024; Lew et al. 2025). Therefore, the literature does not support the idea that Data Mesh is a simple replacement for centralized architectures. It is better understood as a context-dependent approach that redistributes data management responsibilities across the organization.

These findings carry practical implications for managers as well. The literature suggests that Data Mesh is not a default choice but a context-dependent one (Lew et al. 2025). Meaning that it fits best in large enterprises with diverse data domains, whereas smaller organizations may be better served by more centralized approaches. Before adopting it, managers should confirm that several prerequisites are in place: a platform team that can deliver and operate self-serve infrastructure, a federated governance design that balances domain autonomy with shared standards (Dolhopolov et al. 2024), and incentive structures that reward data producers for quality and maintenance, since the team that produces data is often not the team that benefits from it (Baião & Marconi 2024). Clear motivation should itself be treated as a decision criterion, because Bode et al. (2024) warn that organizations sometimes adopt Data Mesh out of hype rather than genuine need. In other words, the relevant managerial question is not whether Data Mesh is

superior in general, but whether an organization has the scale, platform maturity, governance capacity, and incentives to make decentralized data ownership work.

5.2 Academic and Industrial Perspectives

The review also shows differences between academic and industrial literature. Academic papers are stronger in defining and structuring the concept. For example, Machado et al. (2022), Panigrahy et al. (2023), Blohm et al. (2024), Dolhopolov et al. (2024), and Gieß and Hutterer (2025) explain what Data Mesh is, how it differs from related concepts, and which principles belong to it. These papers are useful because they give the topic a clearer theoretical structure.

At the same time, academic literature is still developing. Several papers point out that Data Mesh is a relatively young concept and that empirical evidence is still limited (Winter & Hackl 2023; Bode et al. 2024; Kernstock et al. 2024). Bode et al. (2024) address this problem by conducting interviews with industry experts. Kernstock et al. (2024) also contribute empirical insight through the Alpha case study. Still, the academic literature on Data Mesh remains rather limited.

Industrial literature gives a different but useful perspective. Sources such as Whyte et al. (2022) and Baião and Marconi (2024) show how organizations discuss how Data Mesh is used in practice. Whyte et al. (2022) frames Data Mesh as a response to bottlenecks created by centralized data platforms and teams, and describes on the example of IKEA how this architecture is implemented. Baião and Marconi (2024) gives a more concrete implementation view by describing data products, access approval, metadata, quality checks, and ownership in practice. These sources are useful because they show what Data Mesh looks like outside abstract definitions.

However, industrial sources also need to be used carefully. They often focus on one company, one implementation context, or a consulting perspective. For this reason, their findings should not be generalized too quickly. For example, Baião and Marconi (2024) is valuable as an implementation example, but it does not prove that the same approach would work in every organization. Similarly, Whyte et al. (2022) gives practical insights, but it also reflects a practitioner and consulting-oriented perspective.

All in all, both types of literature are necessary for this topic. Academic literature helps to define the concept and identify research gaps. Industrial literature helps to understand implementation details and practical problems. Since Data Mesh is strongly practice-driven, a review that only considers academic papers would miss important implementation knowledge. At the same time, industrial sources need to be connected back to academic concepts so that the paper remains analytically grounded.

5.3 Research Gaps and Limitations

First, there is limited long-term evidence on Data Mesh outcomes and boundary conditions. Many publications explain the concept or describe early implementation experiences, but there is still little evidence on whether Data Mesh improves data quality, agility, innovation, or business value over time. Future research should therefore examine not only whether Data Mesh works, but also for which types of organizations it is suitable. This is important because the reviewed literature suggests that Data Mesh may be more appropriate for large enterprises with diverse data domains and established platform capabilities than for smaller organizations with simpler data landscapes.

Second, federated governance remains underdeveloped in practice. The literature agrees that Data Mesh needs a balance between domain autonomy and enterprise-wide

standards. However, several papers show that this balance is difficult to achieve in implementation (Bode et al. 2024; Dolhopolov et al. 2024; Kernstock et al. 2024). More research is needed on how organizations can design governance structures, data contracts, metadata processes, and policy enforcement in a way that supports both autonomy and consistency.

This literature review also has limitations. First, the number of analyzed publications is limited, only sixteen of them. Therefore, the concept matrix cannot represent the entire body of Data Mesh literature. Second, the review includes both academic and industrial sources. This was necessary because Data Mesh is strongly influenced by practice, but it also means that the sources differ in scientific quality and purpose. Third, the coding of papers into dimensions involves interpretation. Since many papers discuss architecture, governance, organization, and implementation at the same time, another researcher might classify some papers differently.

Finally, Data Mesh is still a fast-developing topic. New papers and implementation reports may change the state of the literature quickly. For this reason, the findings of this review should be understood as a snapshot of the current discussion rather than as a final assessment of the concept. Collectively, the literature suggests that Data Mesh is promising, but still not fully mature as a research field. Its main challenge is not only technical implementation, but the coordination of decentralized ownership, governance, skills, and platform support.

6 Conclusion

In this literature review I examined the current academic and industry-oriented research on Data Mesh. The findings show that Data Mesh is mainly discussed as a decentralized approach to data management that addresses limitations of centralized architectures through domain ownership, data products, self-serve infrastructure, and federated governance. However, the review also shows that Data Mesh is not a simple technical solution. Its implementation depends on several factors such as organizational maturity, clear responsibilities, governance mechanisms, platform support, and suitable incentives. On the whole, this paper contributes to a structured overview of a still developing research field. Future research should focus especially on empirical evidence, long-term outcomes, and the conditions under which Data Mesh creates real business value.

References

Azeroual, O. and Nacheva, R. (2023) 'Data Mesh for Managing Complex Big Data Landscapes and Enhancing Decision Making in Organizations', in *Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K)*. Science and Technology Publications, pp. 202–212. doi: 10.5220/0012195700003598.

Baião, L. and Marconi, M. (2024) 'From Theory to Practice: Kuehne+Nagel's Data-Mesh Implementation' [Presentation slides]. Available at: https://share.google/Zpt37PSyGyeTLHA01 / (visited on 06/05/2026).

Blohm, I. et al. (2024) 'Data products, data mesh, and data fabric: New paradigm(s) for data and analytics?', *Business & Information Systems Engineering*, 66(5), pp. 643–652. doi: 10.1007/s12599-024-00876-5.

Bode, J. et al. (2024) 'Toward Avoiding the Data Mess: Industry Insights From Data Mesh Implementations', *IEEE Access*, 12, pp. 95402–95416. doi: 10.1109/ACCESS.2024.3417291.

Dehghani, Z. (2019). How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh. Available at: https://martinfowler.com/articles/data-monolith-to-mesh.html

Dolhopolov, A., Castelltort, A. and Laurent, A. (2024) 'Implementing Federated Governance in Data Mesh Architecture', *Future Internet*, 16(4), article 115. doi: 10.3390/fi16040115.

Gieß, A. and Hutterer, A. (2025) 'The Future of Data Management: A Delimitation of Data Platforms, Data Spaces, Data Meshes, and Data Fabrics', *Information Systems and e-Business Management*, 23(4), pp. 971–997. doi: 10.1007/s10257-025-00707-4.

Kernstock, P. et al. (2024) 'Building Industrial Data Platforms — A Case Study on Introducing Data Mesh', in *Proceedings of the 32nd European Conference on Information Systems (ECIS 2024)*. Available at: https://aisel.aisnet.org/ecis2024/track10_dmds_ecosystems/18

Lew, K.L. et al. (2025) 'A Narrative Review of Data Mesh Architecture Principles and Implementation Outcomes', *International Journal on Robotics, Automation and Sciences*, 7(3), pp. 114–123. doi: 10.33093/ijoras.2025.7.3.15.

Machado, I.A., Costa, C. and Santos, M.Y. (2022) 'Data Mesh: Concepts and Principles of a Paradigm Shift in Data Architectures', *Procedia Computer Science*, 196, pp. 263–271. doi: 10.1016/j.procs.2021.12.013.

Pakrashi, A. et al. (2023) 'CowMesh: A data-mesh architecture to unify dairy industry data for prediction and monitoring', *Frontiers in Artificial Intelligence*, 6, article 1209507. doi: 10.3389/frai.2023.1209507.

Palanisamy, G. (2025) 'From Data Lakes to Data Fabric/Mesh: The Future of Enterprise Data Platforms in a Multi-Cloud World', *Journal of Computer Science and Technology Studies*, 7(5), pp. 23–34. doi: 10.32996/jcsts.2025.7.5.4.

Panigrahy, S., Dash, B. and Thatikonda, R. (2023) 'From Data Mess to Data Mesh: Solution for Futuristic Self-Serve Platforms', *IJARCCE*, 12(4), pp. 677–683. doi: 10.17148/IJARCCE.2023.124121.

Rajasekaran, B. and Raj, P.A. (2025) 'Leveraging Data Mesh for Enhanced Data Accessibility and Business Intelligence', *International Journal of Emerging Technologies and Innovative Research*, 12(2), pp. i90–i99. Available at: http://www.jetir.org/papers/JETIR2502809.pdf

Sisson, D. et al. (2024) 'Data Mesh Architecture: Interoperability, Co-Operation, and Co-Regulation'. *SSRN*, 24 July. doi: 10.2139/ssrn.4880709.

Webster, J. and Watson, R.T. (2002) 'Analyzing the past to prepare for the future: Writing a literature review', *MIS Quarterly*, 26(2), pp. xiii–xxiii. http://www.jstor.org/stable/4132319.

Whyte, M. et al. (2022) 'Data Mesh: Just another buzzword or the next-generation enterprise data platform?', *PwC Study 2022: Changing Data Platforms*. Available at: https://www.pwc.de/en/digitale-transformation/data-mesh-the-next-generation-enterprise-data-platform.html / (visited on 06/05/2026).

Winter, R. and Hackl, T. (2023) *Data Mesh at Scale: Exploration of Current Practices in Large Organizations*. University of St. Gallen. Available at: https://www.researchgate.net/publication/379513337_Data_Mesh_at_Scale_-_Exploration_of_current_practices_in_large_organizations

Wolfswinkel, J.F., Furtmueller, E. and Wilderom, C.P., (2013) Using grounded theory as a method for rigorously reviewing literature. *European journal of information systems, 22*(1), pp.45-55. https://doi.org/10.1057/ejis.2011.51

Declaration of Authorship

I hereby declare under oath that I have completed this thesis independently and without using any resources other than those cited; any ideas taken directly or indirectly from external sources, as well as any content created or edited by artificial intelligence such as ChatGPT, are clearly identified as such.

This work has not previously been submitted to any other examination committee in the same or a similar form, nor has it been published.

Hamburg, 09.06.2026

Mariia Bukaieva