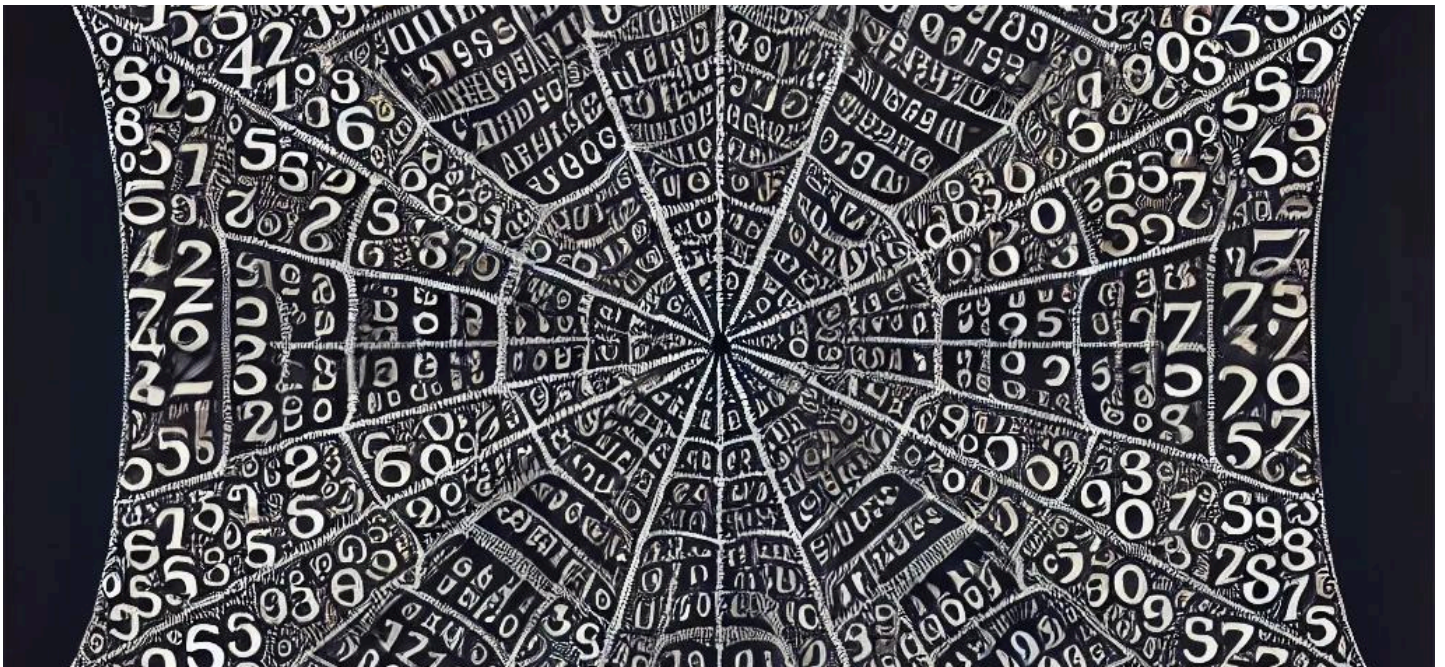




# Data Mesh Architecture (DMA):

A New Solution or Old Idea Repackaged?

By Goldi Tewari Mar 29, 2025 7:30 PM PDT



## Introduction

Organizations frequently encounter challenges in managing and owning data, as conventional methods like Data Lakes and Data Warehouses often fail due to their centralized, hierarchical nature. In response, Data Mesh has emerged as a promising alternative. This innovative framework revolves around key principles: domain-oriented,

decentralized data ownership; the treatment of data as a product; the establishment of self-serve data infrastructures; and the use of federated computational governance.

Notable companies, such as Netflix, JPMorgan Chase, and Zalando, have successfully implemented Data Mesh Architecture, highlighting its potential. By adopting Data Mesh, organizations shift toward an agile, decentralized approach to data management, interlinking various data sources within a "mesh" structure. As a relatively nascent concept, this article critically evaluates Data Mesh and explores its implications for organizations.

## Challenges of Traditional Systems like Data Lakes

**Ambiguity in Ownership and Quality:** Challenges surrounding data ownership and quality frequently undermine the effective management of Data Lakes, adding layers of complexity and high dependence on IT and data governance teams.

**Scalability Constraints:** As the volume of data sources grows, scalability becomes increasingly problematic, often overwhelming teams and infrastructure.

**Cost Efficiency Limitations:** While Data Lakes offer a cost-effective solution for storing vast amounts of diverse, raw data, their inherent shortcomings restrict their potential for broader and more practical applications.

## Evolution of Data Management: From Warehouses to Mesh Architectures

**Data Warehouses and Cloud Tools:** Platforms like Snowflake emerged in the late 1980s. It stores processed and organized data and excels at providing structured, high-performance analytics.

**Data Lakes:** In 2010, data lake came up as a centralized place to store large amounts of data from various sources in its native format. This flexibility of storing raw and unstructured data makes them ideal for big data analytics, machine learning, and real-time processing.

**Databricks and Lakehouse Platform:** Founded in 2013, Databricks serves as a centralized hub for collaborative data engineering and AI. Also known as the Lakehouse Platform, it combines the strengths of Data Lakes and Data Warehouses, enabling real-time data processing, advanced analytics, and seamless collaboration across data teams.

**Delta Lake:** An open-source storage layer introduced in 2019 that enhances Data Lakes with improved reliability, performance, and governance.

**Data Mesh:** Concept introduced in 2019 is best suited for organizations aiming to decentralize data ownership and manage it as an agile product.

# Data Mesh Conceptual Architecture Examples

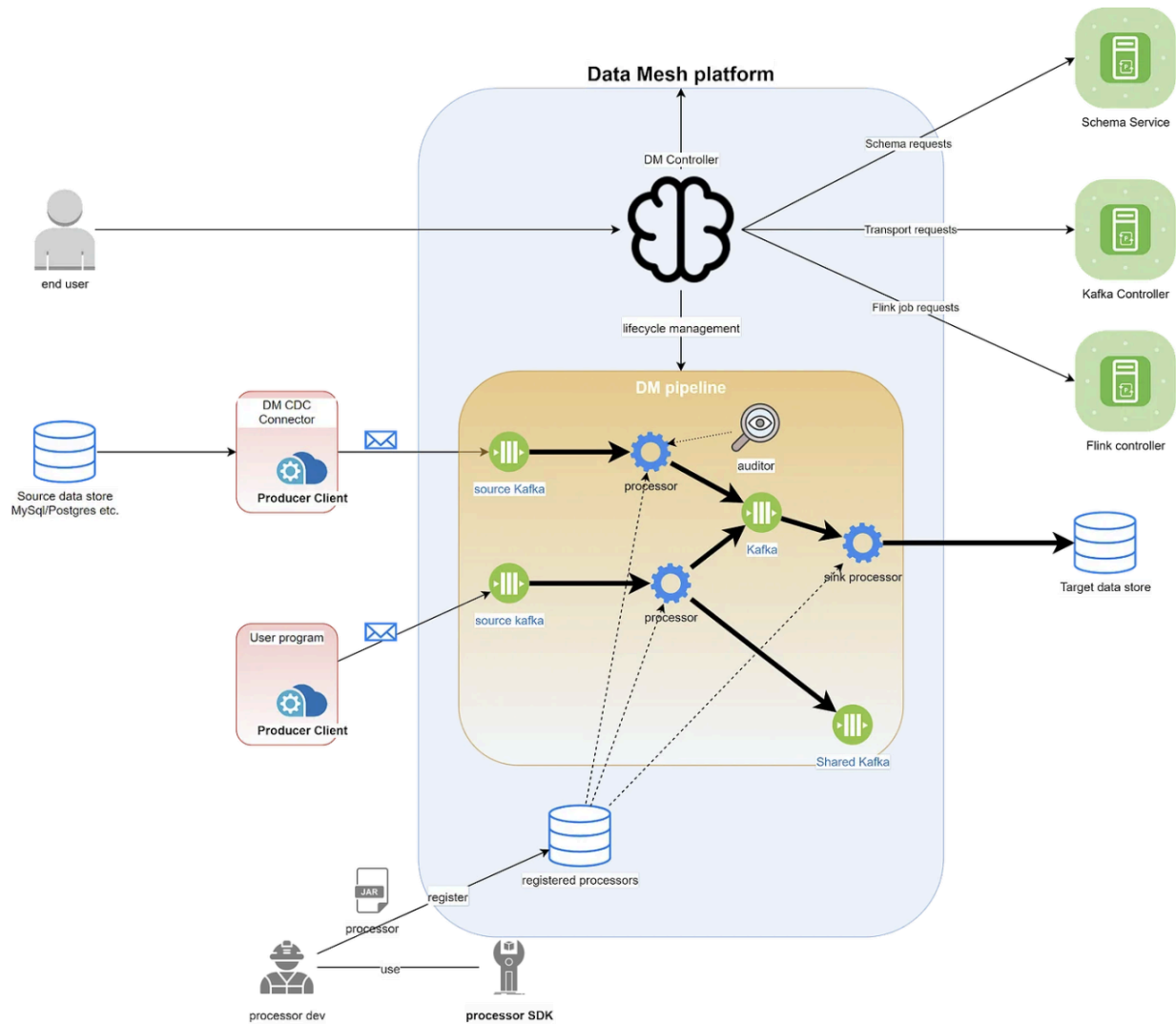


Fig 1. Netflix's Data Mesh Architecture. Credit: [Netflix Blog](#) (2022)

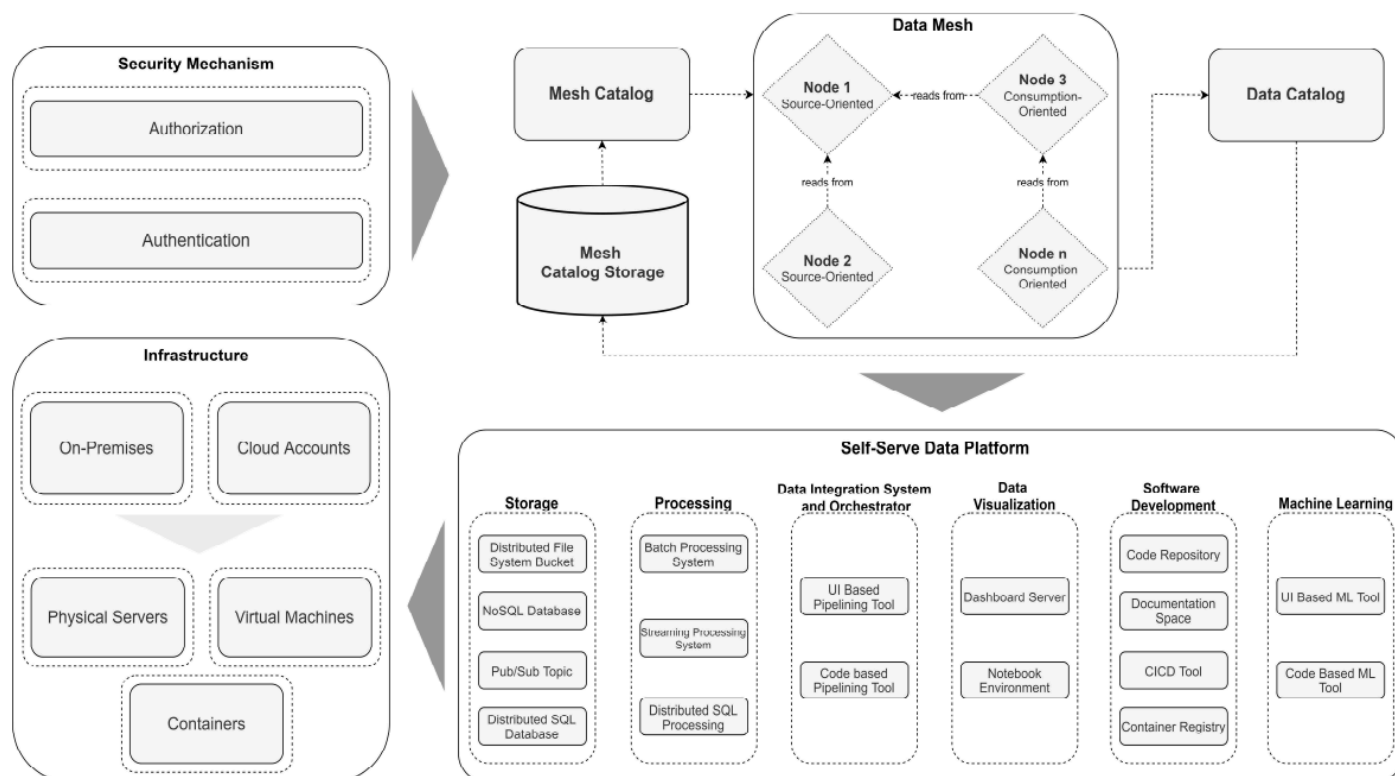


Fig 2. Data Mesh's Conceptual Architecture by Machado, I., Costa, C., & Santos, M. Y. (2021)

## Data as a Product and How Data Mesh Works

Treating data as a product means each data product is agile and consumable by internal and external teams. It is developed, owned, and managed by dedicated teams, ensuring accountability and quality. In this decentralized data system, ownership is distributed among multiple authorities, rather than being controlled by a single entity.

## How Does Data Mesh Work?

In the Data Mesh framework, data is distributed across multiple nodes to eliminate silos. These nodes adhere to four fundamental principles:

1. **Domain-Oriented Ownership:** Teams are responsible for managing and maintaining their data based on their specific domain expertise.
2. **Data as a Product:** Data is treated as a high-quality, user-friendly asset, accessible to both internal and external clients.
3. **Self-Serve Data Infrastructure:** Teams can independently create and manage their own data products, reducing reliance on IT support.
4. **Federated Governance:** While governance standards and rules remain centralized, implementation is decentralized, granting teams the flexibility to adhere to these guidelines.

## Tips for Transitioning to a Data Mesh Architecture for Companies

- Built on the foundation of Flink (as demonstrated by Netflix's Data Mesh Architecture), data mesh processors perform best with Flink SQL. Hence, companies should prioritize adopting Flink SQL before making a switch to Data Mesh. Bo Lei et al. (2022)
- Before adopting data mesh architecture, companies should define distinct domains. These domains must be capable of independently creating and managing their own data products. Additionally, standards for data governance, compliance, security, data exchange, schemas, and APIs should be established in advance, ensuring they operate without reliance on centralized ownership.
- Reliable self-service tools (e.g. Kubernetes, Apache Kafka, AWS Glue, Databricks, Looker, Airbyte, Great Expectations) that support Data Mesh principles should be tried and tested before switching to data mesh.

## Conclusion

Data Mesh represents more than just a rebranding effort and is a progressive evolution of existing data systems. By addressing the limitations of centralized frameworks, it introduces a new paradigm where data is treated as an agile, high-value product. However, organizations should approach this transition as an iterative process, embracing a mindset of continuous improvement to refine and optimize their data organization strategies over time.

## Further Reading

Alademehin, A. (2025). How Data Mesh Architecture is Transforming Traditional Data Governance Models. No, this paper has not appeared or been accepted to appear in any paper series, journal, or book.

Bo Lei et al. (2022). Data Mesh — A Data Movement and Processing Platform Netflix.<https://netflixtechblog.com/data-mesh-a-data-movement-and-processing-platform-netflix-1288bcab2873>.

Chittoor, D. S. (2025). Data Lakehouse Evolution: a Framework for Implementing Domain-Driven Data Mesh Architecture in Enterprise Systems. Technology (IJRCAIT), 8(1), 14-25.

Freno, A. (2017, August). Practical lessons from developing a large-scale recommender system at Zalando. In Proceedings of the eleventh ACM conference on recommender systems (pp. 251-259).

Goedegebuure, A., Kumara, I., Driessen, S., Van Den Heuvel, W. J., Monsieur, G., Tamburri, D. A., & Nucci, D. D. (2024). Data mesh: a systematic gray literature review. ACM Computing Surveys, 57(1), 1-36.

Jean-Georges Perrin and Eric Broda (2024), Implementing data mesh: Design, Build, and Implement Data Contracts, Data Products, and Data Mesh.

Jebasingh, D. (2024). Data Lakes and Data Mesh Architectures: Enabling Scalable and Decentralized Data Governance. *International Journal of Emerging Trends in Computer Science and Information Technology*, 5(4), 16-22.

Katari, A. Case Studies of Data Mesh Adoption in Fintech: Lessons Learned-Present Case Studies of Financial Institutions.

Machado, I. A., Costa, C., & Santos, M. Y. (2022). Data mesh: concepts and principles of a paradigm shift in data architectures. *Procedia Computer Science*, 196, 263-271.

Machado, I., Costa, C., & Santos, M. Y. (2021). Data-driven information systems: the data mesh paradigm shift.

Van den heuvel, J. A. N., Monsieur, G., Tamburri, D. A., & Di Nucci, D. A. R. I. O. (2023). Data Mesh: a Systematic Gray Literature Review.

## Share & follow

**For any inquiries, please contact:**

**Email:** [admin@databananas.com](mailto:admin@databananas.com)

**Website:** [www.databananas.com](http://www.databananas.com)

**YouTube/Instagram:** @databananas