DECENTRALIZING DATA MANAGEMENT IN RETAIL: A REVIEW OF DATA MESH ARCHITECTURE

Ravi Kiran Koppichetti
koppichettiravikiran@gmail.com

Abstract

Data Mesh Architecture is a modern approach to enterprise data management that shifts from centralized systems to a decentralized, domain-focused design. Initially introduced by Zhamak Dehghani, Data Mesh emphasizes four key principles: domain-oriented data ownership, data as a product, self-serve data infrastructure, and federated computational governance. It enables each domain to collect, analyze, and govern data on its own, treating it as a valuable product to improve efficiency.

In retail, Data Mesh empowers units like Sales, Marketing, Inventory, Customer Support, and Supply Chain to own and enhance their data products, supporting data-driven decision-making across the organization at the domain level. Retail domains can ensure higher data quality, timely access, and cross-functional interoperability by decentralizing data management. This paper explores the roles and responsibilities essential to this model alongside the governance and self-serve data platforms that make Data Mesh effective in retail. Ultimately, Data Mesh Architecture provides a scalable, agile framework suited to today's fast-paced, data-rich retail environments.

Keywords: database, data architecture, retail domain, data mesh, data product, data lineage

I.   INTRODUCTION

Managing data well has become essential in today's retail world, where customer preferences shift rapidly, and operations must keep up with constant change. Data Mesh Architecture brings a fresh approach by organizing data around specific business areas—like sales, marketing, and inventory—so each team can take charge of the data that matters most to them. By allowing domain teams to work directly with their data, they can access insights more quickly and utilize them more effectively in their daily tasks. This approach helps eliminate the delays and bottlenecks often in centralized data systems [1]. With Data Mesh, each domain owns its data, fostering accountability and enabling teams to make quicker, more informed decisions. Adopting a Data Mesh approach for retail can mean smoother data sharing across teams and quicker responses to market changes, creating a flexible, customer-centered approach to data that can grow with the business.

II.   DATA MESH ARCHITECTURE

Data Mesh Architecture offers a modern approach to enterprise-level data management and calls for a foundational shift in organizations' architecture, technical solutions, assumptions, and social

structure. It compels us to rethink how we store, manage, and utilize the data to maximize value [1]. In contrast to centralized data architectures, Data Mesh promotes a decentralized, domain-oriented design. This shift empowers individual domains to own and manage their data as a "data product," with designated ownership, quality standards, and interoperability across the organization. In doing so, it addresses vital limitations of centralized data architectures, such as bottlenecks, scalability challenges, and delayed access to data warehouses and data lakes.

Zhamak Dehghani introduced the foundational concept of Data Mesh, outlining four guiding principles: domain-oriented data ownership, data as a product, self-serve data infrastructure, and federated computational governance. These principles help ensure that data is treated as an asset within each business domain, reducing reliance on a central team and allowing quicker, more scalable access to reliable data [1]. As a result, Data Mesh Architecture aims to create a flexible, distributed system that allows for continuous, scalable insights while aligning with business needs.

By applying Data Mesh Architecture, retail organizations can better support data-driven decision-making processes, particularly in environments where agile and scalable data management is critical. In retail organizations, the Data Mesh architecture defines a framework that enables the "data product" ecosystem by establishing organizations and roles, a data governance team and principles, a data platform, data ownership models, and best practices for implementation.

2.1 Domain Oriented Ownership and Domain Teams
In the Retail Industry, Domain Teams genuinely understand how they produce and analyze their domain data, making their role critical. The Data Mesh concept formally acknowledges this expertise.
Domain teams generally manage operational systems, which are often transactional and support both read and write functions. The domain teams themselves are responsible for operating and maintaining these systems. In any organization, only some teams will own their data. Domain-oriented ownership means taking responsibility for owning and maintaining its data models.

2.1.1 Creating Domain Team and designating roles
In Data Mesh Architecture, data ownership is decentralized and structured around specific roles within each Domain, ensuring accountability and quality across the organization. Fig. 1. Shows the Retail Domain, which contains the Retail Operation System of the Domain and
The Key players in data ownership within Data Mesh include:
1. Domain Owner: The key focus of a Domain Owner is Strategic Domain Alignment. A Domain Owner would be someone in Senior Leadership within the Domain who oversees domain strategy, allocates resources, and supports data initiatives.
2. Data Owner: A Data owner focuses heavily on Data Product Strategy and Governance. They would be someone at level of a Domain-Level manager or Lead. A Data Owner should manage the data lifecycle and ensure quality, compliance, and interoperability.
3. Data Steward: The key focus of a Data Steward is operational data management. Individual contributors like Data Practitioners and analysts would be good fits for this role. The Data Steward must ensure data quality, maintain the metadata, and assist domain users with data issues.
4. Data Product Owner: The Product Owner will focus on developing and enhancing data

products based on user feedback. Product Owner or Senior Data Analyst with product management experience is a good fit for the Data Product Owner role in Data Mesh.
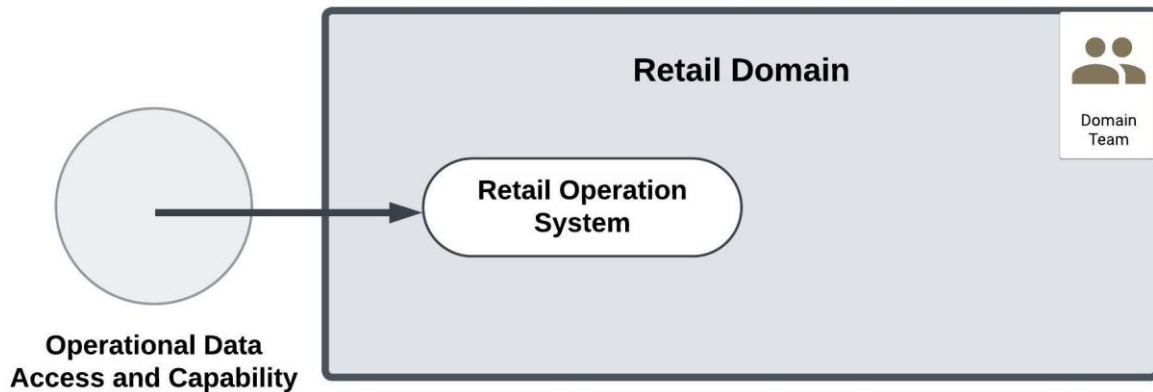

Fig. 1. Domain Team in Data Domain. Adapted from [1]

2.2 Data as a Product

The Data as a Product principle in Data Mesh Architecture shifts the perspective on data, treating it not merely as a by-product of operations but as a valuable asset that should be managed, maintained, and optimized, much like any product a business offers. This principle is particularly impactful in a retail organization, as different units such as Sales, Marketing, Inventory Management, Customer Support, and Supply Chain can enhance their operational effectiveness by managing their data as independent, high-quality "data products."

2.2.1 Implementing "Data as a Product" in Retail Organizational Units

In conventional architectures, data is often stored centrally and accessed as needed by different teams. This approach can create bottlenecks, delays in data delivery, and inaccuracies, which are problematic in fast-moving industries like retail. In the Data Mesh approach, each Domain treats its data as a product. Data is designed, curated, and managed with customer needs in mind. This paradigm ensures that data is accurate, available, secure, and consumable by other units, thus providing the agility and scalability that modern retail organizations need. Fig. 2. visualizes the Data Domain and Data Product.
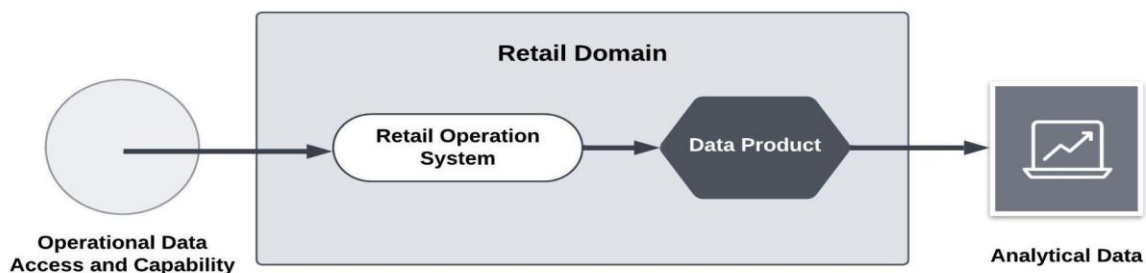

Fig. 2. Data Product in Data Domain. Adapted from [1]

1. Sales Domain: For Sales, owning data on transactions, customer purchase patterns, and sales conversions enables timely access to data that reflects current consumer behaviors. By considering it a distinct "Sales Data Product," the Sales team can respond quickly to market trends, conduct real-time performance evaluations, and better target promotions without depending on a centralized data team [1].

2. Marketing Domain: In Marketing, data on campaign results, customer segmentation, and brand engagement is crucial for optimizing promotional efforts. When the Marketing team manages this as a "Marketing Data Product," they can customize data pipelines, perform rapid analyses, and adjust strategies in real-time. This model supports cross-functional initiatives with units like Sales and Customer Support while meeting governance and quality standards.

3. Inventory Management Domain: Inventory data, which includes stock levels, turnover rates, and demand forecasts, benefits from domain ownership by the Inventory Management unit. This team can make accurate, timely inventory adjustments by managing data as an "Inventory Data Product," reducing overstock and stock outs. Collaborating with the Supply Chain domain becomes more seamless, improving operational efficiency and customer satisfaction

4. Customer Support Domain: Customer Support data, such as feedback, complaint trends, and resolution times, becomes more actionable when owned by the Customer Support team as a "Customer Support Data Product." This ownership model empowers the team to enhance service quality, align responses with customer expectations, and identify common pain points. Additionally, this data is often essential for collaboration with the Marketing and Sales domains

5. Supply Chain Domain: Data on sourcing, logistics, and delivery is critical for the Supply Chain unit. As a "Supply Chain Data Product," this data can be managed directly by the Supply Chain team, enabling faster adjustments to delivery schedules, vendor relationships, and stock planning. Data shared with Inventory Management further enhances inventory accuracy and operational responsiveness

Critical aspects of Data Products Organization are as follows:

1. Direct Domain Knowledge remains within domain teams: All Data transformations are developed and executed by local domain teams to prepare consumable data, eliminating the need for a centralized team to handle data transformations.

2. Decentralized data ingestion: No single, centralized team ingests data from operational systems into a central data platform. For example, Sales data from point-of-sale or Inventory data from Inventory management systems flows directly into domain-specific data products.

3. Data Products function as actual products: Each Data Product has a producer and a consumer within or outside the Data Domain, meeting the needs of at least one specific user. Inventory teams creating stock data for consumption by sales or marketing teams are an example of a Data Product.

4. Data Products are user-centric and read-optimized: Each Data Product is designed and developed with end-user needs in mind, making data readily accessible and optimized for consumption and, for example, optimized sales data for marketing analytics or inventory forecasts for logistics teams.

5. Data Contracts in Data Domain: These contracts serve as a promise to consumers, setting clear expectations for standards, such as data quality and availability. For instance, sales data might include contract specifications on refresh rates and response

times, setting clear expectations for consumer teams in retail.

6. Cross-domain dependencies: Data from one retail Domain, like customer data, can enrich different Domains' data Products, like personalized marketing. Each team manages data within its specific product boundaries.
7. Federated governance within a decentralized structure: Each retail Domain manages its data transformation and pipeline responsibilities while aligning with organization-wide data governance.
8. Data Contract specifications for governance and access: Contracts for each Data Product include critical details like ownership, refresh schedules, data classifications, and access guidelines, ensuring compliance and transparency.

Unlike traditional models where data users directly access systems, Data Mesh allows retail domains to maintain control over their models. This approach reduces the complexity for data consumers, who no longer need to manage operational specifics across systems. The Data Mesh ensures that knowledge about retail domain-specific models stays within the relevant teams, making data more accessible for insights, analytics, and customer experience improvements.

2.3 Federated Data Governance

Federated Data Governance is a set of processes, roles (emphasizing decision rights and accountability), policies, standards, and metrics to optimize business value from data [7]. A Data Governance Team ensures that data products within the Data Mesh Domain meet essential quality, interoperability, and documentation standards. In a retail organization, this team oversees that Data Products align with best practices, are defined through ontologies, are fully documented according to Data Contract standards, and maintain a high level of quality that allows cross-domain interoperability. Although termed a Data Governance team, their responsibilities include elements traditionally associated with Data Management, such as ensuring data quality across the organization. Fig. 3. visualizes Retail Data Domains, Data Products, and Data Governance.

Critical aspects of Federated Data Governance are as follows:

1. Enabling Data Value: Ensure data products are accessible, understandable, and reusable across retail domains, aligning with FAIR (findable, accessible, interoperable, and reusable) data principles [2].
2. Data Contract Compliance: Guarantee that each Data Product meets the terms defined in a Data Contract, which outlines expectations for data accessibility, documentation, and maintenance.
3. Assuring Data Quality: To consistently meet analytical and operational needs, implement and enforce quality standards for all Data Products.
4. Ontology and Model Standards: Develop or provide standardized ontologies for defining Data Product models, ensuring consistency in how data is described and used.
5. Regulatory Compliance: Oversee that the entire Data Mesh architecture aligns with global data governance regulations, including privacy rules for Personally Identifiable Information (PII) and data access controls.
6. Data Stewardship Program: Establish and sustain a data stewardship framework across all retail domains, assigning domain-specific roles to uphold governance standards.
7. Cross-Domain Quality Standards: Define quality metrics and standards that apply across domains, ensuring that Data Products are reliable and usable for all teams in the retail
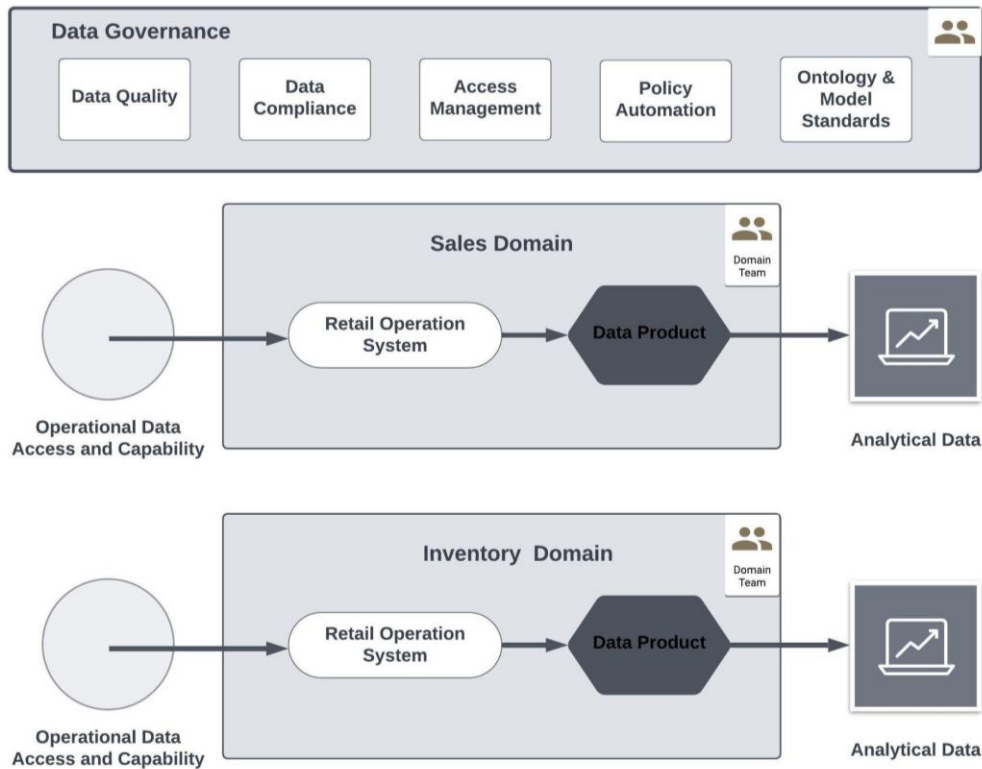
organization.

Fig. 3. Data Product & Governance in Data Domain. Adapted from [1]

2.4 Self-serve Data Platform

The Self-Serve Data Platform in Data Mesh architecture enables domain teams within a retail organization to independently access, manage, and utilize/ analyze data to drive insights and decision-making. This architecture decentralizes data management responsibilities across various teams, such as marketing, supply chain, and sales, by empowering them to build, access, and share Data Products without needing a central data engineering team. Fig. 4. visualizes Retail Data Domains, Data Products, Data Governance, and Data Platforms.

Critical Aspects of Self-Serve Data Platforms are as follows:

1. Data Catalog & Lineage Tracking: A retail Data Catalog allows teams to track customer purchase history, inventory movements, and transaction data across different channels (in-store, online). Lineage tracking is essential for understanding how customer data flows through marketing systems, loyalty programs, and sales forecasting models.
2. Data Storage: Retailers often have vast, distributed data across different geographic regions and systems. A robust data storage platform securely stores all transaction data, customer profiles, and inventory records. The platform must meet industry-specific compliance needs (e.g., GDPR, PCI DSS) while being easily accessible for business intelligence applications.
3. Query Engine: Retail businesses need fast and flexible querying capabilities to analyze sales trends, inventory turnover, and customer behavior. A query engine simplifies access to these critical datasets, allowing business teams to make informed decisions without deep

technical knowledge.

4. Monitoring & Alerting: Retail systems rely on real-time data to manage inventory, monitor sales, and track supply chain performance. An alert system can notify teams of stockouts, system outages, or discrepancies in sales data, ensuring rapid response and minimal disruption.

5. Policy Automation: In retail, policy automation could automate processes like ensuring compliance with data privacy regulations (e.g., customer data anonymization) and restricting access to sensitive financial or customer data based on organizational roles.

6. Audit Trails: In the retail sector, audit trails are crucial for tracking access to customer data, purchase records, and inventory systems, ensuring transparency and accountability for regulatory compliance and internal reviews.
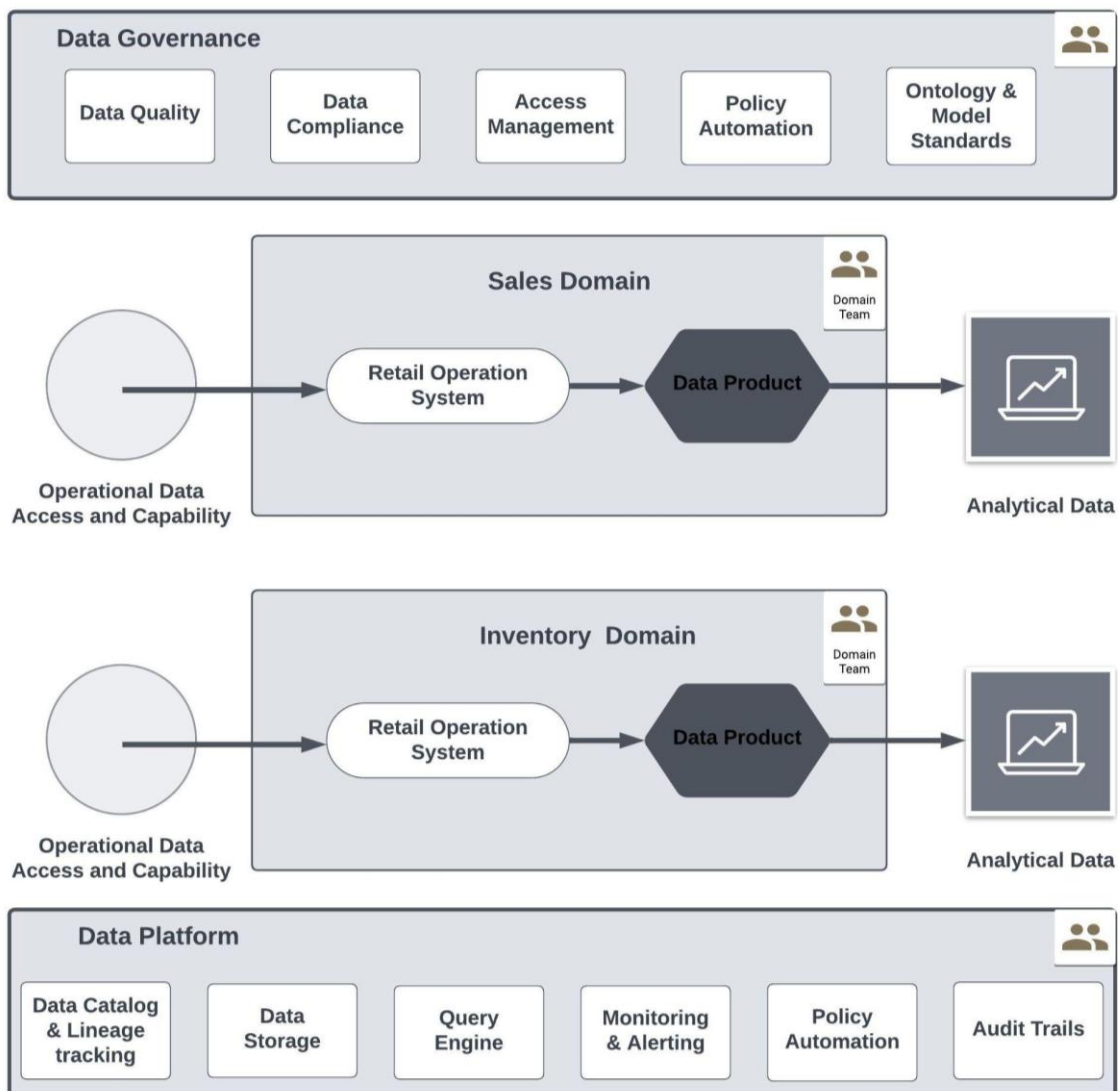

Fig. 4. Data Product, Governance & Platform in Data Domain. Adapted from [1]

III.   ADVANTAGES AND CONCERNS

3.1 Advantages of Data Mesh Architecture in the Retail Industry

Scaling with Data Sources: Since the Data Mesh is a decentralized and domain-oriented design, scaling databases and computational resources is relatively easy, predictable, and can provide high-quality delivery [5]. Data Mesh allows for quick response to market trends, customer demands, and operational changes in the retail industry.

Reduced Bottlenecks: Data Mesh removed the need for a central data team to collect and process data, which reduced bottlenecks and improved overall efficiency in data delivery for analytics. Thus, Retail domains can focus on creating high-quality, reliable data products that meet their specific domain needs [6].

Alignment with Business needs: Data Mesh promotes interoperability and data sharing between domains, enabling teams like sales, inventory management, and marketing to collaborate and derive insights faster. It also enables them to create highly reliable data products, leading to cohesive strategies and improved customer experiences [6].

Enhanced Data Discovery: Data Mesh Architecture encourages domain teams to utilize domain data products effectively. This attribute motivates domain teams to create more self-describing data products and publish product metadata in a central company repository.

Improved Assurance of Compliance: Data Mesh allows companies to distribute governance responsibilities to domains that produce the data. It enables individual domains to create data governance standards and facilitates organizations' compliance with external regulations such as GDPR and CCPA.

3.2 Concerns with Data Mesh Architecture in the Retail Industry

Shortage of skilled Data Engineers: Data Mesh Architecture enables individual domains to collect, process, analyse, govern, and maintain data independently. It increases the need for team members with the required technical skills. The market lacks experienced data engineers, which could affect the implementation and operations of Data Mesh.

Redundancy: Even with a self-serve platform in the Data Mesh, redundancy occurs when building and maintaining data pipelines across all domains.

Change Management: While transitioning to Data Mesh, organizations will face challenges at the technical and organizational levels. Clear communication, employee engagement, and strong leadership to address resistance can increase the chances of successful implementation.

IV.   CONCLUSION

- Data Mesh Architecture is reshaping data management in retail by transitioning from a centralized model to one where each department (e.g., Sales, Marketing, Supply Chain) takes ownership of its data.
- This model enables teams to access, manage, and analyze their data independently, allowing

them to respond more rapidly to evolving customer demands and market trends.

- While the adoption of Data Mesh can pose challenges, such as the need for skilled data engineers and analysts and overcoming resistance to change, it ultimately enhances operational efficiency.
- By decentralizing data ownership, Data Mesh fosters better collaboration and enables retail organizations to provide more data-driven customer experiences.
- Ultimately, Data Mesh empowers retail teams, helping them improve efficiency and agility in a fast-changing business landscape.

REFERENCES

1. Z. Dehghani, Data Mesh Principles and Logical Architecture [Online]. Available: https://martinfowler.com/articles/data-mesh-principles.html. [Accessed: Dec. 7, 2020].
2. L. Corujo, "Who is the FAIRest of them all? Authors, entities, and journals regarding FAIR data principles," Publications, vol. 10, no. 3, p. 31, 2022. doi: https://doi.org/10.3390/publications10030031.
3. A. Machado, C. Costa, and M. Y. Santos, "Data Mesh: Concepts and principles of a paradigm shift in data architectures," Procedia Computer Science, vol. 196, pp. 263–271, 2022. doi: https://doi.org/10.1016/j.procs.2021.12.013.
4. C. Giebler et al., "Leveraging the data lake: Current state and challenges," in Big Data Analytics and Knowledge Discovery, vol. 11708, Springer International Publishing AG, 2019, pp. 179–188. doi: https://doi.org/10.1007/978-3-030-27520-4_13.
5. I. A. Machado, Proposal of an Approach for the Design and Implementation of a Data Mesh (Order No. 30648276). ProQuest One Academic, 2021. [Online]. Available: https://www.proquest.com/dissertations-theses/proposal-approach-design-implementation-data-mesh/docview/2890425511/se-2
6. Z. Dehghani, "How to move beyond a monolithic data lake to a distributed data mesh," Martin Fowler's Blog, p. 45, 2019.