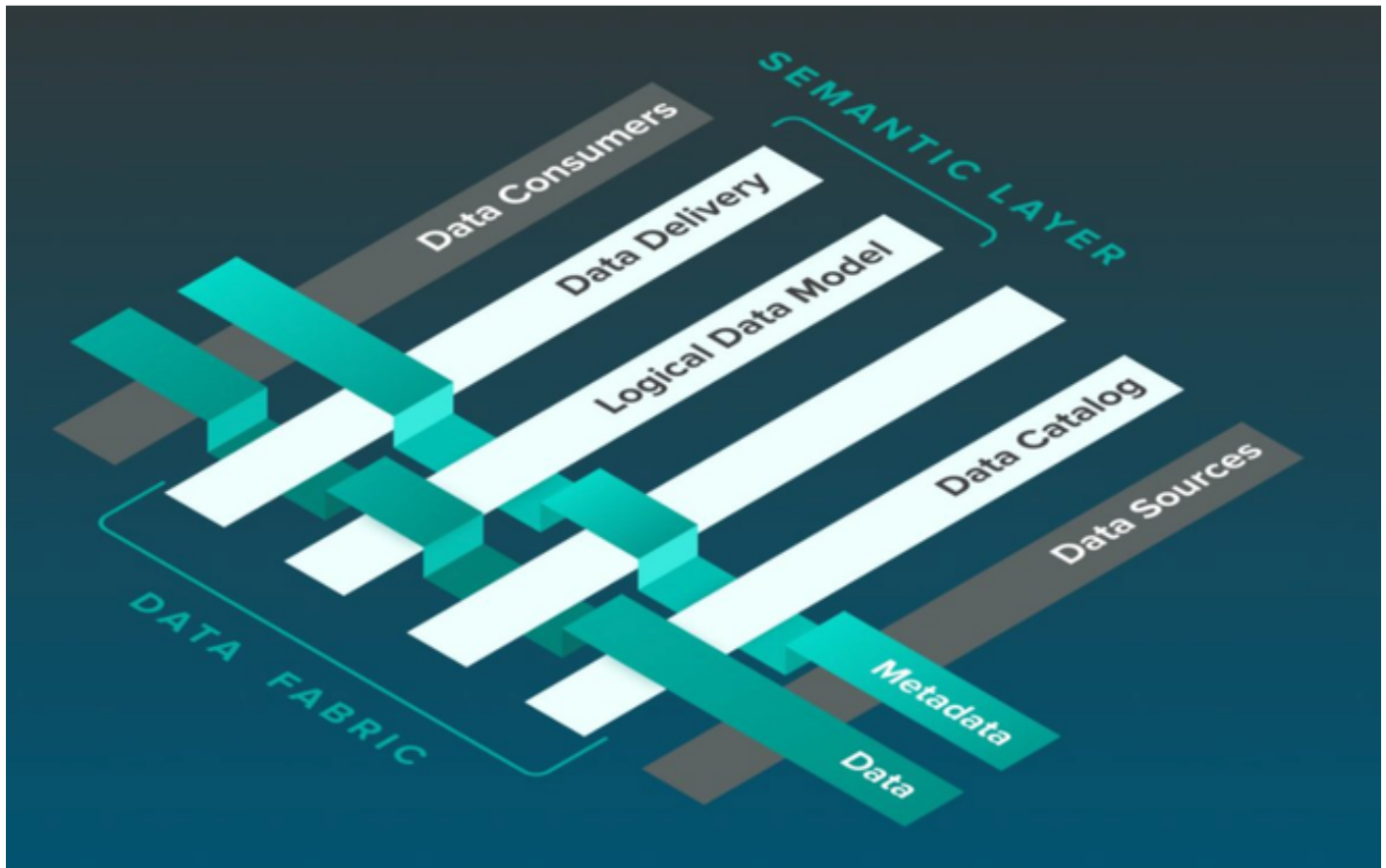# Data Fabric

Jennifer Morales (C)
Contingent Worker

"Data use is exploding across multiple platforms, and organizations desperately need a framework to manage it – to move, secure, prepare, govern, and integrate data into systems. Data fabrics provide that framework, serving as both the translator and the plumbing for data in all its forms, wherever it sits and wherever it needs to go, regardless of whether the data consumer is a human or machine. "

What is a Data Fabric?

Data fabric is an integrated data architectural approach to an enterprise storage operation with the best use of the technologies like cloud, core and IoT devices. It is an architecture that facilitates the end-to-end integration of various data pipelines and cloud environments using intelligent and automated systems.

The most important benefit of a data fabric is that it enables accessing, ingesting and sharing of data in a distributed data environment. It works in a way that it connects to any available data source through the pre-packaged connectors and eliminates the need for programming. It provides data ingestion and integration capabilities.

Data Fabric can be useful in managing multiple environments whether they are on a single premise or multiple cloud networks. It provides built-in data quality, data preparation and data governance capabilities that have been enhanced by

machine learning and augmented automation. It also supports the need for data sharing with internal and external stakeholders through API support.

Data Fabric can help an enterprise to innovate its solution by integrating multiple data sources, compute engines and systems. It helps to reduce downtime and can be used to conduct preventive maintenance analysis. Data fabric show enterprises the needs of their customer enabling organizations to develop strategies to enhance the overall customer experience. Data Fabrics have been impactful in improving the data accessibility for organizations by providing a secure and flexible environment without requiring a massive change to an organization's infrastructure.

A Data Mesh is not a Data Fabric, it is a part of a Data Fabric. It provides an interconnected view into the data. It is a distributed knowledge graph that you can access.

Data Fabric Overview

Data Fabric democratizes data access across the enterprise, at scale. It is a single, unified architecture – with an integrated set of technologies and services, designed to deliver integrated and enriched data – at the right time, in the right method, and to the right data consumer – in support of both operational and analytical workloads.
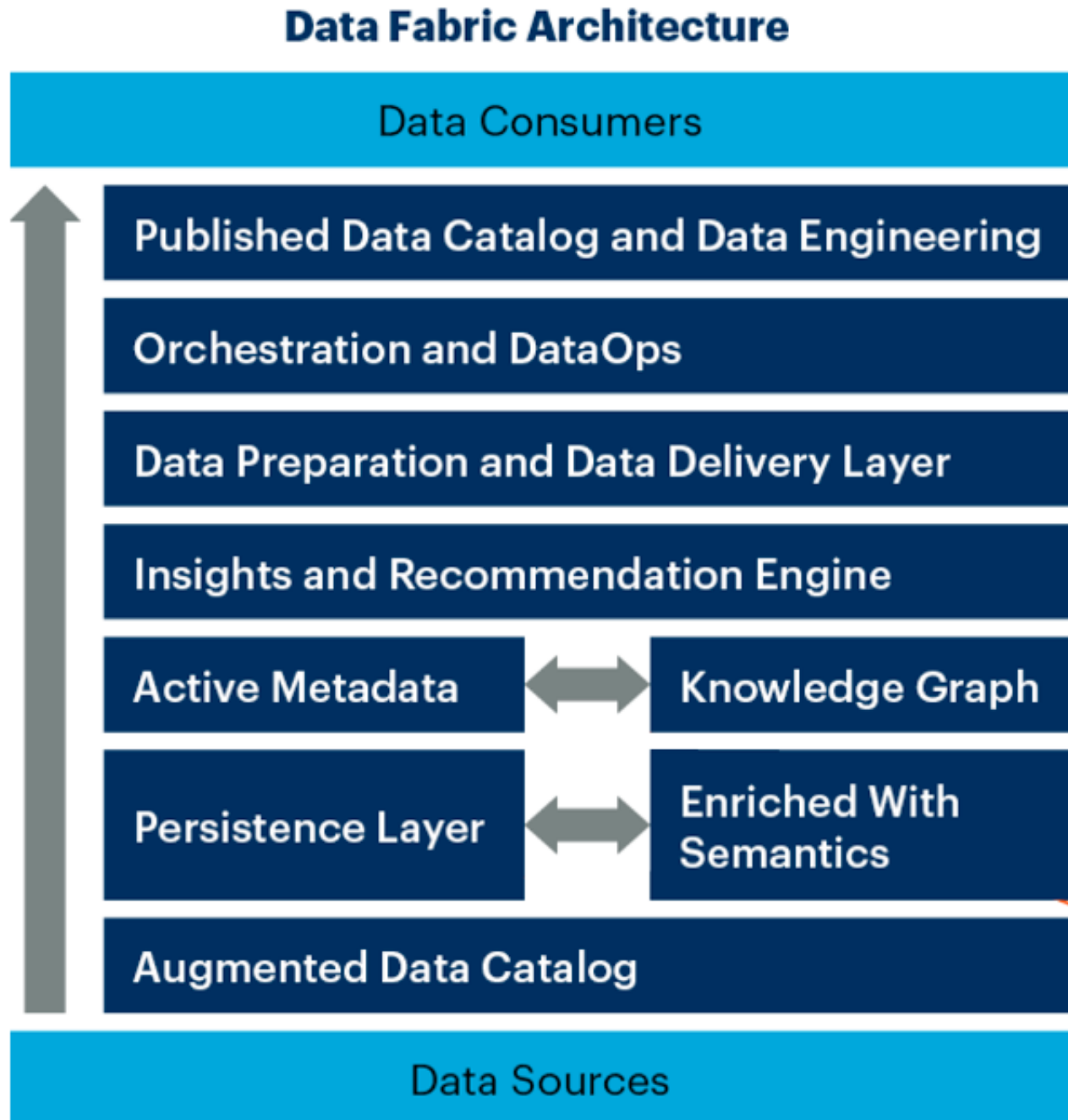
Data fabric combines key data management technologies – such as data catalog, data governance, data integration, data pipelining, and data orchestration.
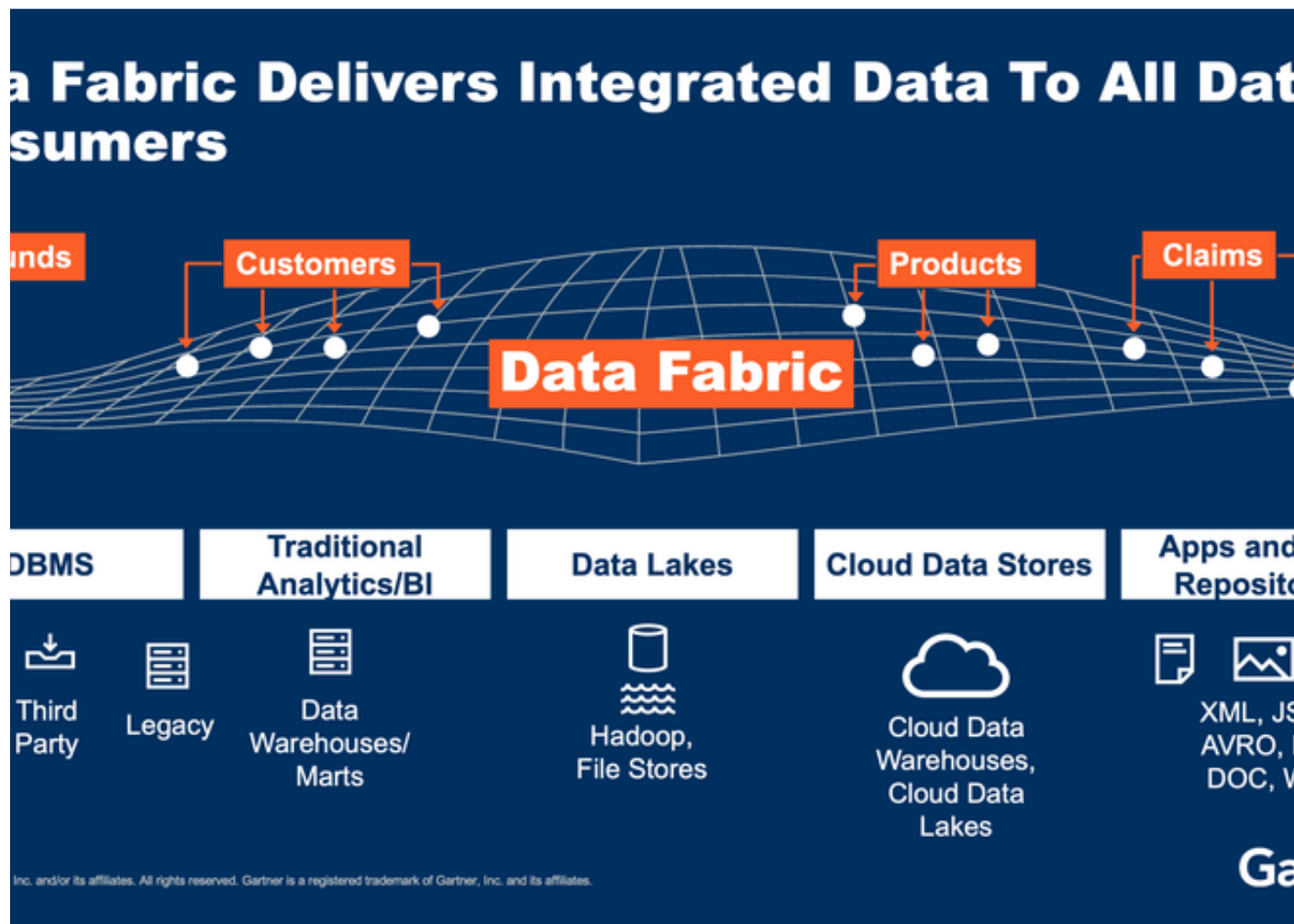
What Is Data Fabric Design?

- A data fabric is a design concept that serves as an integrated layer (fabric) of data and connecting processes. The fabric presents an enterprise wide coverage of data across applications that is not constrained by any single platform or tool restrictions.
- A data fabric follows a metadata-driven approach. Active metadata discovery and semantics inference are key new aspects of a data fabric compared to traditional approaches.
- A data fabric is composable by design. It is made up of components that can be selected and assembled in various combinations.
- Designing a data fabric requires understanding your own maturity as well as the maturity of the various components. We recommend starting with leveraging passive metadata, adapting to knowledge graphs, introducing

active metadata and, finally, planning the orchestration services.

Data Fabric Architecture

Data Consumers
Published Data Catalog and Data Engineering
Orchestration and DataOps
Data Preparation and Data Delivery Layer
Insights and Recommendation Engine

Active Metadata	⟷	Knowledge Graph
Persistence Layer	⟷	Enriched With Semantics

Augmented Data Catalog

Data Sources

As the diagram above illustrates, as data is provisioned from sources to consumers, it is cataloged, enriched to provide insights and recommendations,

prepared, delivered, orchestrated, and engineered.

Data Fabric Core Capabilities

Key Capabilities

Data fabric supports the following key capabilities integrated into a single platform:

1. **Data catalog**
   a. To classify and inventory data assets, and represent information supply chains visually
2. **Data engineering**
   a. To build reliable and robust data pipelines for both operational and analytical use cases
3. **Data governance**
   a. To assure quality, comply with privacy regulations, and make data available – safely and at scale
4. **Data preparation and orchestration**
   a. To define the data flows from source to target, including the sequence of steps for data cleansing, transformation, masking, enrichment, and validation
5. **Data integration and delivery**
   a. To retrieve data from any source and deliver it to any target, in any method: ETL (bulk), messaging, CDC, virtualization, and APIs
6. **Data persistence layer**
   a. To persist data dynamically in a broad range of relational and non-relational models

Non-functional Capabilities

Data fabric should also address the following key non-functional capabilities:

1. **Data scale, volume, and performance**
   a. Dynamically scale both up and down, seamlessly, no matter how large the data volume.

  b. Support both operational and analytical workloads, at enterprise scale.

2. **Accessibility**
   a. Support all data access modes, data sources, and data types, and integrate master and transactional data, at rest, or in motion.
   b. Ingest and unify data from on-premise and on-cloud systems, in any format – structured or unstructured.
   c. The data fabric logical access layer needs to allow for data consumption, regardless of where, or how, the data is stored, or distributed – so no in-depth knowledge of underlying data sources is necessary.

3. **Distribution**
   a. Data fabric should be deployable in a multi-cloud, on premise, or hybrid environments.
   b. To maintain transactional integrity and data governance capabilities, data fabric needs to support a smart data virtualization strategy.

4. **Security**
   a. Where data is persisted, it must be encrypted and masked to meet data privacy regulations.
   b. Data fabric should be able to deliver user credentials to the source systems, so that access rights are properly checked and authorized.

Source articles for this page:
- [Data Fabric Architecture is Key to Modernizing Data Management and Integration](#)
- [What is a Data Fabric? | AtScale](#)
- [Using Data Fabric Architecture to Modernize Data Integration](#)
- https://insidebigdata.com/2022/03/26/the-framework-that-will-connect-tomorrows-it-is-edge-data-fabric/
- https://www.k2view.com/what-is-data-fabric
- [Data Mesh](#)
- [https://towardsdatascience.com/what-is-a-data-mesh-and-how-not-to-mesh-it-up-210710bb41e0](#)

- [What is Data Mesh? | Starburst](#)
- https://futurice.com/blog/when-should-organizations-consider-data-mesh
- [Intuit's Data Mesh Strategy](#)
- [What is a data fabric? | IBM](#)
- [What Is a Data Fabric? | NetApp](#)
- [What is a Data Fabric?](#)
- [Six Top Data Fabric Use Cases](#)

Ideation & Emerging Technology   Edit   **Contact Us**