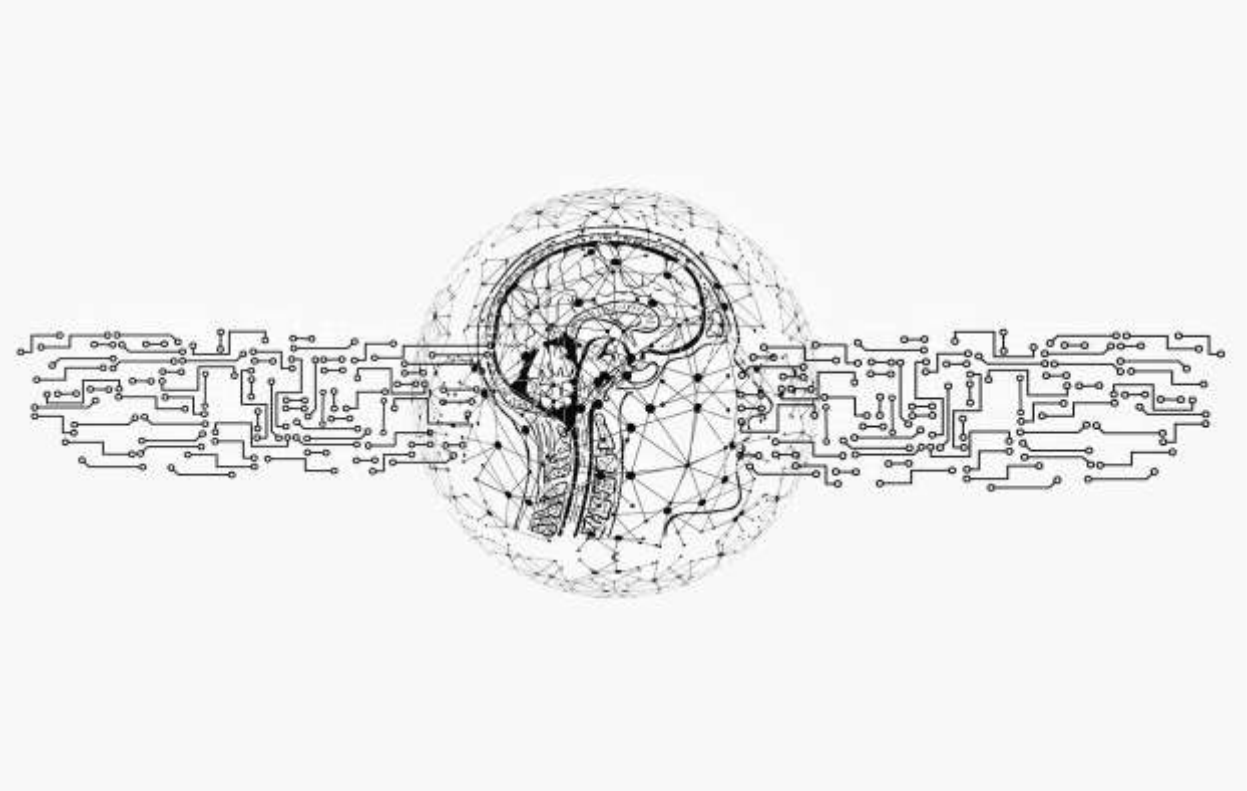
The Data Fabric as a Value Generation Engine

- A Strategic Framework for AI Readiness, Real-Time Decisions, and Unified Governance

By John Santaferraro, CEO and Head Research Analyst of Ferraro Consulting



Published by FERRARO Consulting

www.ferraroconsulting.com

Sponsored by Cloudera, Inc.

www.cloudera.com

The Data Fabric as a Value Generation Engine

- A Strategic Framework for AI Readiness, Real-Time Decisions, and Unified Governance

Three major technology trends are rapidly changing the requirements for information technology modernization: the unification of everything, generative AI, and agentic AI. First, we are seeing the unification of multiple data stores into a single store, multiple data engineering tools into a single platform, and multiple analytical approaches into a unified analytics layer. Second, generative AI has radically changed the way people engage with data at every stage of the data lifecycle, from data generation to insight. We have moved from visual to verbal interaction with data. Third, agentic AI has moved us from simple automation to autonomous agents that perform the work of the various roles throughout the entire data lifecycle operating in complex business ecosystems.

What is a data fabric? How is it different from a data mesh? A data fabric is a data management architecture that provides unified access to diverse data for the entire data management lifecycle. All data engineering and analytics operate in a fully integrated and completely interoperable manner. The data fabric operates as one, providing physical and logical access to a broad range of sources and insight delivery to many different targets, including machines, applications, and devices. The most advanced data fabric architectures actively use metadata for AI-enablement, automations, copilots, and visualizations across the entire fabric.

- The data fabric is not a data mesh. A data mesh is a paradigm focused on delivering self-service data products to business units, putting the business fully in control of data, analytics, and AI. It focuses more on people and processes than on technology.
- The data fabric is not a product; it is an architecture. There is not one single vendor or product that provides a data fabric. However, it is possible to measure vendors based on their alignment with the architecture.

Both Gartner and Forrester have well-defined data fabric definitions and architectures, along with vendor evaluation. The Ferraro Consulting definition aligns with both industry analyst firms, and this paper provides additional drill-down detail on the functionality and interoperability encompassing the broad market offerings.

Why does a data fabric matter? Three things drive the need for organizations to adopt data fabric technologies. One, AI is driving organizational data flow beyond previous levels; and the previous levels were already outpacing data teams' ability to meet business demands. Two, the effort to deploy, integrate, and maintain a disintegrated data management practice is significantly greater than the resources required to acquire and operate a data fabric. Three, the emergence of AI along with the proliferation of real time data flows shrinks the window within which companies need to make decisions.

What are the criteria for selecting a data fabric? Because data fabrics cover a broad spectrum of data engineering and data management capabilities, it requires a compelling starting point and well-designed expansion strategy. There are several factors that impact data fabric selection: vendor offerings, your existing technology landscape, your existing technology skillset, your business use case, internal readiness, future direction, and technology investment planning.

From the perspective of technology evaluation and buying criteria it is vital to understand which elements of the data fabric are foundational and which are ancillary. Buyers should select vendors based not just on immediate requirements but also based on the potential for adding services as new capabilities are needed in the future. For this reason, vendors who cover the broad range of data management and analytics capabilities rise above those with limited services or point solutions.

At its core, the data fabric is built on the active use of metadata. Select a vendor whose metadata strategy enables multiple services actively using a single, integrated set of metadata, making the data fabric highly extensible as technology continues to advance. For example, if your primary need is to streamline data pipelines, do not buy a data pipeline tool, buy a data pipeline platform that is part of a data fabric, with a rich set of metadata at its core.

In addition, if your organization is expanding its use of AI, metadata services like data governance become essential investments. Especially for AI, look for vendors that unify the governance of both structured and semi-structured data. Since both types of data are necessary for generative AI and agentic AI, this is not an option. The goal of AI is to increase both automation and autonomy, both of which require excellence in data quality to better ensure actions taken by the AI.

What are the necessary basic components?

Because of the way a data fabric unifies access, data engineering, and analytics, it consists of five distinct layers. Each layer contains a set of stores or services that all interact with each other within their common layer. They also interact with services in other layers. This matrixed interoperability, connected in all directions by metadata and metadata services, forms the fabric concept of the data fabric. The five layers are AI Services, Analytical Services, Metadata Services, Data Services, and Data Stores. The data stores are not part of the data fabric, but the provision of data stores by a data fabric vendor helps unify analytics.

LAYER ONE: Metadata Services

Metadata lies at the core of the data fabric. Without the curation of a unified, rich set of metadata within the data fabric, it is not possible to operationalize AI and analytics in a safe and reliable manner. Technical metadata is necessary to understand the data and to track its lineage as it moves through the various phases of data engineering and the various layers of the data fabric. Business metadata is necessary to connect the entire data lifecycle, from data generation to insight delivery, bringing context and insight to the entire business.

Metadata services rely on the active use of a rich and common set of metadata to knit together the fabric, but most importantly, to activate the use of data, insight, and AI throughout the enterprise. The following metadata services often have overlap from one service to another, especially in the area of governance. Mature data fabric vendors rationalize the necessary services for the data fabric and eliminate the reduplication of effort present in multi-platform environments.

Data Catalog Services

Data catalog services actively use metadata to help organizations easily find and understand data, understand data usage, and actively use this data intelligence for development of both operational and analytical applications. It is the first place users of all types go to begin their work with data. Think of the data catalog as the intelligent heart of the data fabric. It is not just a passive inventory of data; it is an active metadata system of record for all data, intelligence, and AI. To operate as the heart of the business, the data catalog typically includes a business glossary, technical metadata, operational metadata, active data lineage, conversational metadata, a search interface, plus AI-driven discovery, classification, and generation of metadata.

Data Governance Services

Data Governance services balance centralized standards with distributed domain ownership. IT and governance teams serve as "expert guides," providing the infrastructure and policy templates that enable data product owners within business units to define and enforce their own data standards. AI enablement makes data governance dynamic, protecting the organization from risk while guaranteeing the accuracy of domain-specific decision-making.

AI Governance Services

In the context of a data fabric, AI governance is the operational system that translates business objectives, regulatory obligations, and risk requirements into measurable technical controls and defensible evidence. It ensures that AI models and AI-driven applications can be deployed, monitored, and scaled responsibly, without slowing innovation or relying on one-time reviews. Unlike policy-driven or checklist-based approaches, AI governance in a data fabric is continuous, automated, and deeply integrated into data and AI pipelines.

Data Engineering Automations

Data engineering automations are continuously running services that use metadata as active intelligence to optimize data processing environments. These automations combine static rules, policies, and heuristic algorithms to activate years of accumulated data management and data engineering best practices. Within the data fabric, metadata becomes prescriptive and operational. Typical automations include active metadata creation and collection, semantic layer creation, performance optimization, automated data pipeline creation, automated data pipeline fix, cost optimization recommendations, schema tracking, data drift corrections, data quality, and many more.

Generative AI Assistants

In the context of a metadata-centric data fabric, generative AI assistants translate natural language questions and goals into fabric-aware action, and they guide humans through the fabric based on historical usage and tribal knowledge. They use large language models combined with continuously updated metadata to interpret intent, provide contextual guidance, and generate recommendations, explanations, and actions across the fabric. Unlike static interfaces or task-specific assistants, copilots operate across platforms, domains, and lifecycle stages.

Autonomous AI Agents

Autonomous AI agents are continuously operating services that perceive, reason, and act on behalf of humans to operate the data fabric. Unlike automations which run based on rules and machine learning, or AI assistants which take step actions, autonomous AI agents use a basic level of perceiving, reasoning, and probability to take action according to their programmed guardrails and objectives. For intelligence and autonomy, they utilize metadata to provide the context, limits, and feedback required for safe and effective operation. Agents transform the data fabric from a responsive system into an adaptive, self-directed operational environment, while remaining constrained by governance, intent, human-defined boundaries, and humans in the loop.

LAYER TWO: Data Stores

The data fabric provides unified access and governance to diverse data without exposing the complexity of multiple data platforms to users of all types. Much of the pipeline design and data engineering is done at the metadata services layer, but modern data platforms or data stores also do an increasing amount of data processing with compute for different types of workloads. The types of data supported by the data fabric include structured data, unstructured data (or semi-structured), streaming data, edge data, visual data, and audio data. The most common data store configurations today are the data

warehouse, data lake, data lakehouse, object store, and operational database. There is an increasing trend toward multi-model and multimodal support in a single platform, as well as a move toward universal access where all of the processing is done at the data fabric layer and access to data is virtualized; both in support of zero-copy approaches. While data stores are not a necessary part of the data fabric, there is an advantage to having both the data fabric and the data stores in a single, unified environment.

LAYER THREE: Data Services

Data services are services that touch data in any way throughout the data lifecycle, from creation to action. Within the data fabric they operate within a well-governed, cataloged, automated, AI-enabled, unified environment. While these data services mimic some of those described in the metadata services section of the data fabric, much of the work can also be done on the data platform, where the data is stored.

Streaming

Data streaming services collect data moving at high speed, enabling real time curation and analysis of data streams, as well as support for event-driven applications and AI. Within the data fabric, data in motion is curated along with other types of data and deployed along with them in singular data pipelines. As organizations modernize their architectures, we are seeing a shift-left movement in which more curation, transformation, and analysis is performed earlier and at lower cost while data is still in motion.

Edge Streaming

Edge streaming services capture, process, and route data directly from devices, sensors, and distributed environments, enabling real-time transformation and movement of high-velocity data as it is created. With edge data as a first-class citizen in the enterprise dataverse, data engineers build pipelines that also include streaming data, structured, and semi-structured data with full lineage, governance, and interoperability.

Discovery

Data discovery services enable data analysts and data engineers to understand the structure and condition of their data assets early in the data engineering lifecycle. Over the last several years, we have seen data discovery move from specialized tools to visualization and natural language discovery provided by the data platform and data fabric vendors at the data store level.

Pipeline

Data pipeline services enable data analysts and data engineers to design and deploy data pipelines, either in a low code - no code environment, or with manual coding or coding tools. Cross platform and cross application pipelines reduce integration costs and speed time to decision.

Cleansing

Data cleansing services enable data analysts and data engineers to transform data to achieve consistency and accuracy for the purpose of insight delivery, AI enablement, and operational embedding. Data quality requirements and processing remain different for structured and semi-structured data. However, with rich metadata, generative AI, and agentic AI, we are seeing unification of data cleansing at both the data services and metadata services levels.

Integration

Data integration services provide the necessary tools to bring together different data sources and different types of data into a singular data pipeline terminating in a singular storage of data. Within the data fabric, integration spans the full spectrum of all data assets, applications, APIs, and data streams.

Transformation

Data transformation services provide data analysts, data engineers, and business analysts with a standard set of tools to unify data for use in analytical, AI, and operational use cases, by applying business logic, semantics alignment, and structural changes.

Ingestion

Data ingestion services collect and move data from a wide range of systems, formats, and sources (both batch and streaming) into the various data stores in a consistent, governed, and metadata-aware way. Within the data fabric, they ensure that all incoming data is immediately discoverable, trusted, and ready to interoperate with other core data services.

Sharing

Data sharing services provide the well-defined, governed, secure, capabilities needed to safely share data across the extended business ecosystem of partners, customers, regulators, and internal teams. Within the data fabric, they ensure that shared data is consistently discoverable, trusted, and interoperable so every other service in the fabric can use it safely and appropriately.

Products

Data product services take data sharing to the next level, enabling the sharing of carefully curated and governed data for specific parts of the business, such as human resources, marketing, campaign management, or supply chain. Within the data fabric, data products include curated datasets, data streams, intelligence, APIs, or ready to use AI.

Observability

Data observability services allow data teams to see, monitor, and manage the health and usage of their data across the entire hybrid cloud: on-premises, public cloud, and at the edge. Within the data fabric, observability covers the entire data landscape, unites all data assets in a single view, and is increasingly automated as AI-enablement matures. Mature observability goes beyond data to provide workload and infrastructure for performance and reliability.

LAYER FOUR: Analytical Services

The analytical layer for the data fabric is complicated due to the rising need for access to both data and curated insight via visualization, analytical, and natural language interfaces, as well as the interface to machines, devices, models, and now agents. While it is not necessary for the data fabric to provide all of these analytical services, more mature platforms have unified access and analytical services and offer them as part of or along with their data fabric platforms.

Visualization

Visualization services provide interactive, intuitive ways for everyone, from data engineers and analysts to business users, to explore, understand, and communicate insights at every stage of the data engineering and analytics lifecycle. Within the data fabric, visualization becomes metadata-aware, allowing users to access governed metrics, semantic definitions, lineage, and health status directly within or accessible to their visual environment.

Analytics - Visual and Technical

Analytics services unify drag-and-drop, low- or no-code visual analytics with powerful coding interfaces in a single, collaborative environment where business users and technical experts can work side-by-side using the tools and interfaces that suit them best. Within the data fabric, analytics are powered by shared metadata, governed metrics, and consistent semantics, ensuring that visual insights and technical analyses draw from the same trusted foundation and can be easily combined, reused, and operationalized.

Natural Language

Natural language gives users the ability to ask questions, explore data, and generate insights using everyday language, while also leveraging the data fabric's metadata to understand business context, intent, and meaning. Advanced language models play the role of a data analyst, business analyst, or business professional. Within the data fabric, natural language evolves into agentic analytics where agents can answer questions, suggest next steps, highlight anomalies, and guide users toward deeper insights.

MLOps

MLOps services manage the full lifecycle of machine learning, from ideation and experimentation to deployment, monitoring, and continuous improvement. MLOps requires different capabilities to manage more complex computing and to ensure that models remain reliable, governed, and aligned with business objectives. Within the data fabric, MLOPs services increase value when they have interoperability with a broad range of data services and AI services for unified pipeline management, optimization, and automation.

LAYER FIVE: AI Services

AIOps

AIOps services require a different approach than analytics or MLOps. The training, containing, and use of both language models and AI agents require well-defined and structured business semantics, as well as tight coupling between business and technical metadata to ensure accuracy. In addition, while AI agents are not yet prolific, the data fabric architecture must be ready for metadata-driven orchestration for multi-agent environments. Within the data fabric, generative AI and agentic AI rely on business semantics, metadata, and continuous AI governance to ensure accuracy, safety, and alignment with business values and vision. Advanced systems translate regulatory obligations and risk requirements into measurable technical assessments and defensible evidence across the entire AI lifecycle.

Generative AI

Generative AI services combine model-centric capabilities like LLMs, fine-tuning, embeddings, and retrieval, with business-centric output like content generation, research, and even high-value strategic use cases such as acting as an executive coach or transforming a consulting firm's knowledge base into a client-facing product.

Agentic AI

Agentic AI services introduce true agency, designing and deploying intelligent software programs that can mimic how humans accomplish tasks, fill roles, collaborate in departments, and operate across entire ecosystems, embodying the core attributes of intelligence, autonomy, action, reactivity, interactivity, adaptability, and clearly defined limits.

Enterprise Requirements

In order to meet the needs of the enterprise, the data fabric must include the functionality necessary for consistent, constant, resilient, agile, and safe operations.

Unified Control Plane

A single control plane orchestrates the entire data fabric for a unified operational experience. Control should cover deployment, resource allocation, and lifecycle maintenance. This unified approach eliminates operational silos, ensuring that the fabric remains manageable regardless of the activation of new capabilities, new locations, or the operationalization of new datasets, analytical applications, or AI agents.

Centralized Security and Zero-Trust Enforcement

Centralized security consolidates identity management, encryption protocols, and access policies into a single, unified enforcement framework. Because of the constant interchange of data and data pipelines across diverse types of data, different layers of architecture, and across multi-cloud and hybrid locations, centralized security is the only way to maintain consistent security for all data.

Data Privacy and Sovereign Autonomy

The need for adherence to local governmental regulations and for the protection of intellectual property are necessary for the data fabric, especially because of its use of data from across the organization, and in many cases, across country lines. The data fabric needs to run in both private and sovereign scenarios, without the loss of functionality or control. The combination of "localized" governance rules and a globalized metadata view maintains privacy and sovereignty while enabling an enterprise-wide view of the data.

Hybrid and Multi-Cloud Consistency

Hybrid, cloud, and on-premises consistency is a top requirement for a data fabric platform. In the best case scenario, the singular data fabric contains the same functionality, operates the same way, and is controlled by a single control module whether data resides on-premises (local servers or data centers), in the cloud (public, private, or multi-cloud setups), or in hybrid configurations that combine both. Consistency eliminates the expense and confusion created when multiple instances of the “same” software are marked by unexpected inconsistency. Portability and interoperability give the enterprise the flexibility it needs, especially to run a global business ecosystem.

Elastic Enterprise Scale

Scale for the modern data fabric must be both horizontal and vertical to accommodate growth of data volumes, diverse data types, differing data latencies, and increasing user concurrency. An enterprise-grade fabric manages this expansion elastically across disparate geographies and compute environments without a degradation in service or manual intervention. As we rapidly move toward the expansion of AI, the elastic backbone must be able to support thousands of concurrent data products and AI agents with reasonable stability. Mature data fabric vendors already have plans in place for cost control and economy of scale.

Balanced Performance

For the enterprise, performance must be multi-faceted. The enterprise workload mix includes high volume transactions, real time execution, and high performing compute for complex analytical workloads, often spread out in highly distributed landscapes. Advanced data fabrics meet more granular performance requirements for query execution and data movement, using intelligent optimization, caching, and well-orchestrated distributed computing. This type of multi-faceted performance enables high-fidelity for real-time decisions, automations, and autonomous AI.

Functional Breadth

The breadth of a vendor's offering determines the fabric's ability to support the complete data-to-AI lifecycle without relying on brittle integrations between disjointed point solutions. An enterprise-grade platform provides a comprehensive suite of native, interoperable capabilities—including advanced ingestion, semantic modeling, data observability, and AI orchestration—all within a single ecosystem. This breadth reduces technical debt and eliminates the need for "Frankenstein" architectures, ensuring that every phase of the data lifecycle adheres to the same unified standards and governance protocols as the organization evolves from simple reporting to sophisticated AI workflows.

The Data Fabric as a Value Generation Engine

Like other modernization initiatives, data fabric return on investment is measured not just by cost savings. The data fabric has the potential of generating value in seven ways: AI and generative AI readiness, unified data engineering productivity gains, real time decision intelligence, unified governance and regulatory responsiveness, democratized self-service, technical debt reduction, and innovation acceleration.

Generative AI and Agentic AI Readiness

The number one barrier to entry for AI remains data quality, for both structured and unstructured data. The data fabric goes beyond data quality to provide AI with clean, semantically rich, context aware data via its rich metadata foundation. Mature data fabrics also provide rich streams of data that serve as triggers for both the Large Language Models (LLMs) and AI agents. By automating the data supply chain for AI, enterprises can move models from prototype to production in weeks rather than months, creating a competitive multiplier that drives market leadership.

Unified Data Engineering Productivity

The average large enterprise has ten different data engineering and metadata services platforms. By unifying these platforms with a data fabric architecture, end-to-end data pipelines are quicker to deploy, more reliable for operations, and require far less time for repairs. Value is created by eliminating the "integration tax"—the endless hours engineers spend moving data from one platform to another and fixing broken data flows. For the enterprise, this translates into a 60-80% decrease in IT support tickets and allows highly skilled talent to shift from maintenance to high-value innovation.

Real-Time Decision Intelligence

Streaming data intelligence used to be required only for niche use cases, but now that almost every modern company has a digital presence, it is no longer an option. With the onset of generative and agentic AI, real time triggers become even more important for autonomous decisions. With the event-driven data fabric, real time triggers ignite actions all across the enterprise, with timely precision, ensured accuracy, and minimized risk.

Unified Governance & Regulatory Responsiveness

Unified governance in a data fabric becomes an active, background process rather than a resource-intensive manual effort. It ensures that data across all business units at all stages of the data lifecycle adheres to local, regional, and global regulations (like GDPR or CPRA). Universal coverage makes response to regulatory reporting and inquiries almost automatic. Value is created through risk-reduction, especially when the fabric prevents costly data breaches and regulatory fines before they occur. Ultimately, well-governed data and AI give executives a level of trust that would be extremely difficult without the data fabric.

Democratized Self-Service

A rich set of both technical and business metadata enables quick search for data and insight along with all of the contextual information necessary to make informed decisions. Within a data fabric, well governed data products enable business users to find data quickly. In addition, support for natural language questions helps users understand the intricacies of insight, as well as expand their knowledge to make more accurate decisions.

Technical Debt Reduction

Modernization, especially with data fabric architecture, embraces existing technology while laying a technology foundation for the future. The data fabric allows enterprises to wrap their old legacy systems in a modern interface, extracting value from legacy data without the risk of a "rip-and-replace" migration. Value is created by extending the life of existing investments while providing a low-risk bridge to a fully cloud-native future.

Ferraro Consulting POV on Cloudera

Ferraro Consulting sees Cloudera as a leader in the data fabric market due to the work they have done over the past several years to integrate their platform and add the necessary data management and data engineering capabilities necessary for an integrated, interoperable, singularly managed platform. Most importantly, the Cloudera investment in a central metadata store and recent investments in metadata services companies lays the foundation for flexibility and adaptability as AI takes shape and the data landscape continues to morph.

With regard to the data fabric, Cloudera has three strong differentiators, along with numerous products across the entire data fabric architecture.

1. **Consistent, active metadata glues the Cloudera data fabric together.**
   Cloudera uses metadata consistently across all analytics, data, data management, and data services. In addition, with Trino they govern access to the data. Therefore, they also have consistency in security and access. In addition to SDX, their core metadata services platform, they are rapidly integrating the data lineage and other metadata services from their Octopai AI acquisition.

2. **All of the Cloudera data fabric software is available on any platform, anywhere.**
   There are many different data sources and targets being accessed by the data fabric. Cloudera brings it together by offering all their software components on any infrastructure. What Cloudera formerly marketed as “true hybrid,” is now the “anywhere cloud” capability with their acquisition of Taikun. Data and software can be on AWS, Azure, Google Cloud, plus the data center. Any other cloud will also be easy to add given the Taikun acquisition. They can simply deploy Kubernetes layers wherever the need arises.

3. **The Cloudera platform supports all analytics and all types of data.**
   Cloudera provides the complete range of data fabric capabilities for any analytics engine and for all types of data. The customer can start with a data lake or data warehouse, start with batch or streaming. They can even bring all the data into an object store, then put Iceberg on top of it.

Cloudera Data Fabric Alignment

The Cloudera Unified Data Fabric offering includes or is augmented by the following, complete set of offerings. The breadth of Cloudera’s data fabric and data platform offerings is an advantage for buyers wanting to grow their data, analytics, and AI capabilities.

AI Services

The Cloudera AI Services offering is excellent, with capabilities that AI-enable the platform, as well as cover the full range of AI capabilities including generative AI and agentic AI. The combination of Cloudera AI Inference Services with NVIDIA provides full scalability and elasticity for AI agents, assistants, and applications. Cloudera AI Studio enables private generative AI and agentic workflows at different skill levels. Cloudera AI Assistants are available for visual analytics, discovery, and all SQL-based activities via a natural language interface, as well as AI Copilot for the full end-to-end pipeline lifecycle.

Analytical Services

Cloudera Analytical Services are strong. From analytical data pipelines to visualization, from streaming data analysis to MLOps, Cloudera provides analytical services for data scientists, data engineers, data analysts, business analysts, and business users all on a unified data platform. Cloudera supports analytical capabilities throughout the entire data lifecycle. Cloudera Streaming enables Flink analytics on data in motion. Cloudera Data Engineering supports analytics as part of the data pipeline, running analytics in a database or data lakehouse for optimal efficiency. Cloudera Data Visualization adds natural language discovery and analytics, as well as automated dashboard capabilities. Cloudera MLOps includes operational capabilities along with AMPs, Accelerators for ML Projects.

Metadata Services

Cloudera Metadata Services are very strong. Cloudera's active metadata strategy has matured from the original vision of knitting together different open source data management tools, to its current vision of unifying, governing, and securing the entire data and AI lifecycle. Cloudera SDX supports a full range of data catalog, governance, and security in a single shared experience across the entire Cloudera portfolio. The recent acquisition of Octopai, now known as Cloudera Data Lineage, augments the Cloudera offering with automatic metadata generation and the active use of data lineage for context-aware analytics.

Data Services

Cloudera Data Services are very strong. Cloudera Data Flow enables the collection, curation, and analysis of streaming data. Cloudera Streaming operationalizes Kafka and Flink for the full management of data in motion for the enterprise. Cloudera Data Engineering orchestrates, operationalizes, and automates complex data pipelines across the entire data lifecycle. Cloudera Edge Management deploys NiFi for intelligent control of edge devices with real time edge data collection and management. All of these data services run on Cloudera's unified data platform with unified access and interoperability with their data warehouse, data lake, data lakehouse, and operational databases, enabling both MLOps and AIOps. Cloudera Observability goes beyond data observability to monitor and manage the health and usage of data, infrastructure, and workloads across the entire hybrid cloud: on-premises, public cloud, and at the edge.

Data Stores

Cloudera Data Store offerings are very strong. Cloudera provides a data warehouse, data lake, data lakehouse, operational database, and object storage options to round out their standing as a single-source provider of a unified platform with full breadth of functionality. In addition, the Cloudera platform can leverage a host of third-party data stores across the complete range of infrastructures, allowing organizations to leverage their existing investments.

Enterprise

Cloudera Unified Data Fabric is excellent and fully enterprise ready. What Cloudera originally did to harden the Hadoop environment for enterprise use; it has now done for its entire unified platform. Cloudera provides a unified management console; however, they do not deploy management and governance in a SaaS environment. Their entire portfolio, on premises and cloud, functions with elastic scaling, even allowing "cloud bursting" to supplement data center requirements. The company does provide centralized security and zero-trust enforcement across all products. Cloudera is unique in the way it supports hybrid and multi-cloud versions of all offerings with complete consistency. With their investment in Taikun, their offerings become even more flexible.

For more information, contact Cloudera at info@cloudera.com, or visit www.cloudera.com.