



Embracing a Holistic Approach:

Data Virtuality's Mission for Data Integration and Management
Inspired by Data OS Principles

Emerging from our first-hand experience of data integration challenges, Data Virtuality has consistently pursued a comprehensive and integrative approach to making data accessible and usable. We always strive to provide a holistic data integration and management approach that abstracts away the technical complexity to empower data users. Our vision is to enable use cases with the optimal set of data management and integration features.

In light of growing complexity and the expanding user base over the past decades, we believe it's more crucial than ever to embrace a philosophy of data integration and management that resembles an operating system - rather than merely a middleware or platform.

Drawing inspiration from the concept of an operating system, we aim to emulate a 'Data OS' experience for our users. This focus is reflected in the following key aspects:



Flexible and scalable data integration: Data Virtuality incorporates multiple integration styles in a single platform, including data virtualization, ETL/ELT, streaming, and CDC to enable the most efficient solution for different use cases.



Effortless data access: Data Virtuality provides a unified point of access for all users, ensuring that everyone technical or non-technical can efficiently access and utilize data.



Universal language: Data Virtuality introduces a single language for data querying and manipulation, streamlining the data integration process and reducing the specialist bottleneck and learning curve for users.



Adapt to the cloud-centric landscape: Data Virtuality offers various deployment options: SaaS, self-hosted in the cloud, hybrid, and on-premises.

By emphasizing these aspects, we aspire to offer a trusted solution for organizations seeking to benefit from more holistic and flexible data integration and management practices.

● Navigating the specialized world for universal data accessibility and utilization

Despite two decades of rapid pace of digitalization and data-driven initiatives, many companies still struggle to become insight-driven or data-informed. Current approaches to solving data integration and management issues, often focusing on the technology by acting as a middleware, have inherent limitations, leading to increased complexity and cost.

To empower data-driven decision making, four critical challenges centered around data accessibility and usability must be overcome:

Challenge 1:

Managing an ever-growing number of data sources and tools

The constantly expanding pool of data sources has led to the development of numerous data tools, resulting in a surge of complexity and costs. This is compounded by the existence of various data types and formats that must be integrated to yield usable data. The sheer array of options makes it challenging for organizations to identify the most suitable combination of data solutions that align with their specific needs and objectives.¹

Challenge 2:

Reducing manual efforts in data accessibility for analytics

Organizations often grapple with the challenge of transforming raw data into actionable insights. This involves cleaning, organizing, processing, and delivering data to make it ready for analysis. However, many current data management tools require the involvement of specialists that are hard to find, leading to high manual efforts, bottlenecks, and valuable insights remaining buried within the data.²

Challenge 3:

Enabling a broader spectrum of users and use cases

Today's workforce is evolving, with an increasing number of business users building and managing technology and analytical capabilities. These "business technologists" require flexible, intuitive tools that allow them to leverage data in their daily operations and decision-making processes.³

Challenge 4:

Streamlining and unifying processes around the data.

The growing complexity around the data necessitates new approaches to organizing the data-related processes. This is essential for improving collaboration, increasing agility and speed, and establishing a culture of continuous improvement. These issues fall under the umbrella of what we call DataOps.⁴

¹<https://www.bcg.com/publications/2023/new-data-architectures-can-help-manage-data-costs-and-complexity>

²More than 90 percent of those surveyed said that it was challenging to some degree to make data available in a format usable for analytics. <https://www.matillion.com/resources/blog/matillion-and-idg-survey-data-growth-is-real-and-3-other-key-findings>

³According to Gartner research, 41% of employees can be described as business technologists.

<https://www.gartner.com/en/articles/the-rise-of-business-technologists>

⁴What is DataOps? <https://www.rackspace.com/solve/what-dataops>

Enter the World of Data OS

In addressing the challenges of data accessibility, usability, DataOps capabilities, and the rise of business technologists, the Data OS has the potential to revolutionize how organizations and individuals leverage data to make informed decisions.

Like an operating system, a Data OS standardizes the processes that make data accessible and usable, abstracting away the peculiarities of each source, providing a single point of access, and a single language to interact with all sources and applications

This streamlined approach radically simplifies the process of making data workable for reports and data applications. By embracing the Data OS concept, Data Virtuality enables the modern workforce from the average business user to the experienced data professional to harness the power of data.

● Drawing Parallels:

How the operating system analogy fits to the Data Virtuality concept

The current state of data integration and management shares striking similarities with the early days of operating systems in the 1960s. Back then, operating a computer required specialized knowledge of machine code and manual entry of instructions through switches on the machine's front panel. Debugging involved examining I/O ports and restarting the process from scratch, making it a tedious and highly specialized task. This complexity hindered mass adoption of computing technology.

The introduction of Windows revolutionized the operating system landscape by managing hardware resources and providing a single access point and language for running applications. This made personal computing accessible to a broader audience, as users no longer needed to be experts in machine code to interact with their computers.

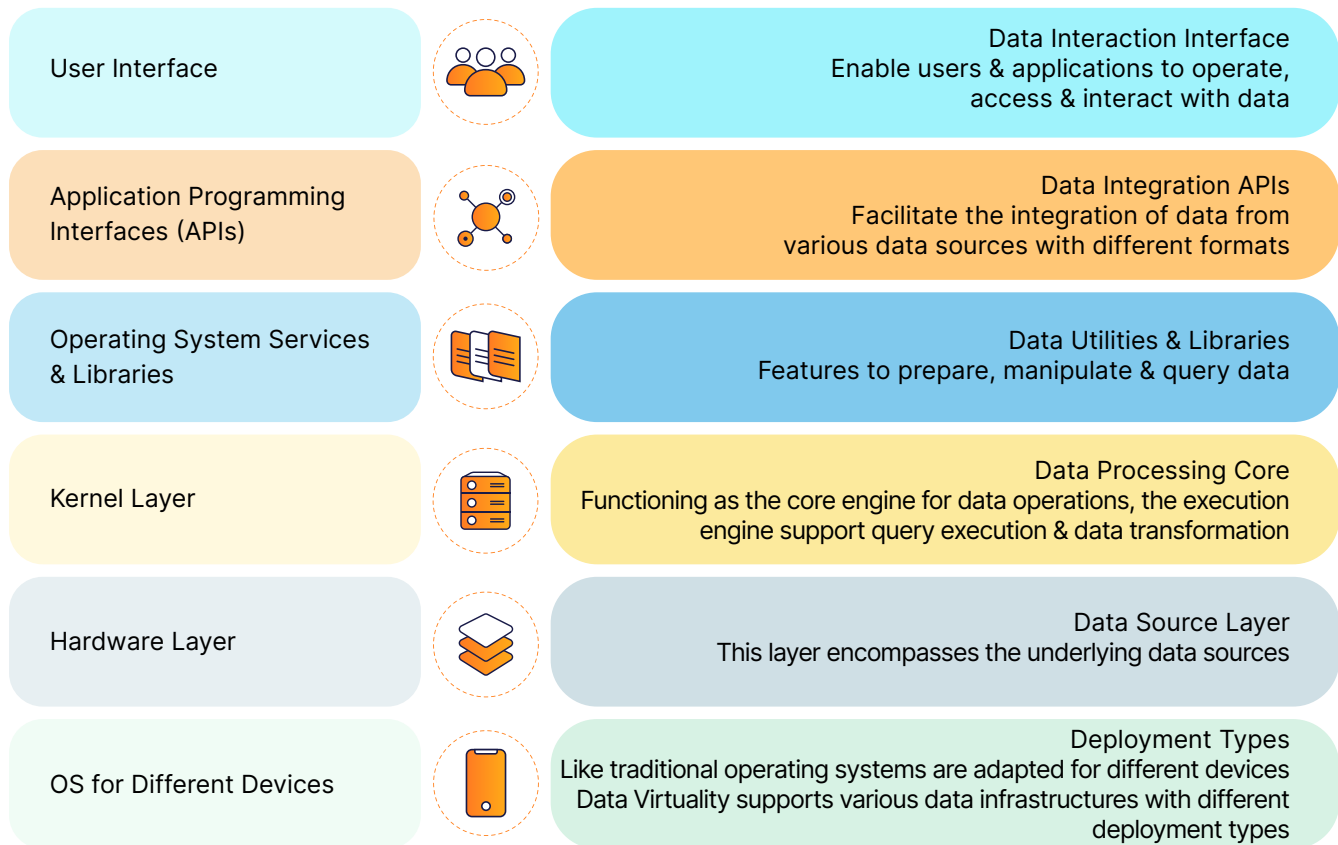
Similarly, data integration and management have long been the domain of specialized experts. A multitude of data formats, platforms, and technologies often requires specialized knowledge and skills to navigate and extract value from the data. This inherent complexity creates bottlenecks due to the limited number of experts available and slows down the overall process of data-driven decision-making.

The concept of Data OS addresses these challenges of complexity and bottlenecks in data integration and management in the same way that modern operating systems revolutionized personal computing.

A Data OS is not only composed of different layers similar to an operating system, but it also offers a range of AI-driven automation features to minimize errors in manual processes. Specifically, in the realm of DataOps, the Data OS provides features like Git integration, data quality automation, data integration as code, and CI/CD support. These features assist in tasks such as provisioning, monitoring, scheduling, and version control. The Data OS offers uniform programmatic access to these features, enabling easy integration into broader DataOps processes within an organization.

Traditional OS

Data OS



● Benefits of a Data OS

A Data OS provides several benefits that help organizations improve their data management and analysis capabilities. **These include:**

Unified Data Access:

A Data OS provides a single point of access to all data sources, abstracting away the peculiarities of each source and enabling seamless integration across different systems, platforms, and formats.

Enhanced Data Usability:

By making data more accessible and organized, a Data OS empowers users to derive actionable insights more easily. It helps transform raw data into a usable format for analytics, making it more valuable for decision-making processes.

User-Friendly Interface:

Designed to be intuitive and user-friendly, a Data OS caters to the needs of both technical and non-technical users. This enables a wider range of employees to leverage data in their daily operations and decision-making processes.

Scalability and Flexibility:

A Data OS is built to scale with the evolving data needs of modern organizations. It can adapt to changing requirements, ensuring smooth performance and minimizing bottlenecks as data volumes and complexity grow.

Improved Data Governance:

A Data OS helps organizations implement robust data governance practices by providing transparency, traceability, and control over data assets. This ensures data quality, security, and compliance with relevant regulations.

Rudimentary Analytics and Visualization:

Data OS often includes native essentials like rudimentary analytics and visualization capabilities, enabling users to analyze data and gain insights. This further supports data-driven decision making within the organization.

By addressing these aspects, a Data OS empowers organizations to become truly data-informed, optimizing data-driven decision making and unlocking the full potential of their data assets.

● Data Virtuality vs query accelerators:

integration vs connection

Navigating the data integration and management solution market can be complex due to similar marketing messages from various technology vendors. When comparing Data Virtuality with query accelerators, one must understand their fundamentally distinct perspectives and focal points.

Data Virtuality offers a data integration and management solution that essentially operates as a Data OS. Its primary focus is on integrating data from a multitude of disparate sources, making them accessible and usable, irrespective of the storage system. These data integration features include critical operations such as data movement, data replication, change data capture (CDC), and extract-transform-load (ETL) capabilities.

In contrast, query acceleration vendors concentrate on enhancing data querying and retrieval speeds. Their focus is primarily on optimizing data access and retrieval in data lake environments. Moreover, some query accelerators even have the capability to delegate queries to connected data sources. However, despite their ability to connect to external data sources, they lack the comprehensive functionality to integrate the data from these sources, since they do not contain the data integration features mentioned above.

The absence of such features necessitates the maintenance of additional tools, leading to several challenges:

- Increased software costs due to the requirement of maintaining multiple tools.
- Additional burden of software maintenance due to the increased number of tools.
- Disruptions in the architecture, causing incomplete data lineage that requires manual maintenance.
- Lower productivity levels as users need to familiarize themselves with different sets of tools.

In light of these points, organizations should meticulously assess their specific needs and use cases when deciding between these distinct solutions. The choice between integration (as offered by Data Virtuality) and connection (as provided by query accelerators) is not a trivial one and warrants careful consideration.

● Data OS vs Data Mesh / Data Fabric:

Complementary elements in the data management ecosystem

Data OS cannot be compared to Data Mesh or Data Fabric, as they each serve different roles within a data management strategy. While a Data OS is a software component, Data Mesh⁵ and Data Fabric represent data architecture designs or concepts. With its focus on the users and use cases, Data OS can well serve as a foundation of a Data Mesh or Data Fabric, providing the necessary tools for data management. This illustrates that Data OS, Data Mesh, and Data Fabric are not competing concepts but rather work in concert to create a robust data management ecosystem that meets the needs of modern organizations. This crucial understanding helps in discerning the unique roles these elements play in a comprehensive data management strategy.

Mission:

Enable organizations and all their people to make better use of their data

Data integration and management are vital in the ever-expanding digital landscape. Our Data Virtuality Platform offers a comprehensive solution resembling a Data OS, addressing accessibility, usability, DataOps capabilities, and the rise of business technologists. By drawing parallels with the evolution of operating systems, Data Virtuality envisions a simplified, unified approach to data integration and management.

⁵Data mesh is a decentralized sociotechnical approach to share, access, and manage analytical data in complex and large-scale environments



● About Data Virtuality

“The Data Virtuality Platform is a flexible tool that allowed us to enable several different use cases such as data governance, data quality checks, self-service BI, data preparation for regulatory reporting, digital marketing analysis, and refinement of the implementation of the business logic.”

Data Architect, Banking (taken from Gartner Peer Insights)



B/S/H/



WELEDA
Since 1921

- Acknowledgements:
Honorable Mention in 2022 Gartner Magic Quadrant for Data Integration Tools

- Awards:
Most Innovative Data Management Provider 2022, 2021, and 2019 (A-Team Insights),
2020 and 2019 Deloitte Technology Fast 50

Request Demo: demo@datavirtuality.com

Free Trial: <https://eu.platform.datavirtuality.com/#/start-trial>

Message: info@datavirtuality.com

Visit: datavirtuality.com

