The Case for Data Governance within Data Analytics and Data Integration Initiatives

Data governance has become one of the most strategic and impactful data disciplines to be adopted in organizations across the globe. It is particularly relevant to organizations investing in data integration, warehousing, and analytics capabilities. This paper shares the benefits of a mature data governance strategy and how Data Virtuality has aligned its approach to supporting our clients' varied data governance needs.

Data governance adoption has helped to address many historical challenges that companies have grappled with in the past, namely:

- Lack of engagement for enterprise data policies and processes
- Insufficient knowledge of business terms and data definitions
- Inadequate data access and permission management
- Wasted cost of repeated data quality improvement
- Poor lineage and provenance of data sources
- Lack of data ownership and accountability

When ignored, these challenges can cause data initiatives to take longer, cost far more than expected, and deliver an inferior experience for internal and external end-users/customers. Nowhere are these impacts more keenly felt than during the build and deployment of a data warehouse and analytics solution.

Benefits of Data Governance for Data Integration and Data Analytics

Trusted decision-making: The insights from data analytics are only as good as the raw data supplied. Poor quality data can create mistrust of data warehousing and reporting solutions, hampering aspirations to adopt a data-driven style of decision-making. Adopting data governance helps assure the quality and trust of the data flowing into the data analytics environment.

Faster defect resolution: Whenever data defects impact the analytics service, a broad-ranging apparatus should be in place to help coordinate with the upstream teams to locate and quickly remove the source of the problem (e.g. leveraging data lineage). Data governance ensures that this end-to-end assurance approach is actively managed and supported at all levels.

Increased data democratization and accessibility: One of the goals of data governance is to drive greater value by exposing more data to a wider community across the organization. Data governance plays a crucial role in coordinating data accessibility and security controls through data owners and stewards – critical for enabling enterprise data democratization, particularly as emerging trends such as data mesh call for greater federation of governance, interoperability and domain-oriented data ownership.

Faster delivery: When organizations have a detailed understanding of the meaning, structure, quality and lineage of their data, they can dramatically increase the speed and productivity of the data warehousing build process. If they've purchased a pre-built data warehouse, data governance helps them to understand the data more easily and share that knowledge with other teams who are actively building reports and feeds that connect to the warehouse.

Comprehensive reporting: Data governance provides a framework for creating accurate and commonly agreed data definitions and business terms. Through these kinds of metadata, the data analytics function can more easily identify the information required by the business and ensure easy accessibility within the reporting and analytics environment.

> Note: While it is useful to implement data governance measurements using standalone tools, it's important to bear in mind that standalone tools are focused on governing the metadata only. A more successful approach is to implement data governance in the context of an overall data management strategy that combines metadata and data in real-time to fulfil the governance aspects.

How can the Data Virtuality Platform Support and Enhance your Data Governance Goals?

Our mission is to help businesses leverage the full potential of their data through a single source of truth platform that connects and manages all data. Therefore, operationalizing data governance outcomes, such as deploying a common accessibility, controls, and standards framework for data, has long been a familiar goal within the Data Virtuality solutions roadmap.

Before we explain how data governance is enabled in our platform, it's helpful to understand our products' overall architecture because our unique approach lends itself naturally to the operationalization of data governance and data quality.

Introducing the Data Virtuality Approach to Data Integration

The Data Virtuality Platform combines two distinct technologies to create a unique approach to data integration. The combination of data virtualization and next-generation ETL enables a data infrastructure with unparalleled flexibility and performance.

Data Virtuality offers data teams the freedom to select the ideal integration approach for each use case instead of being locked into a restrictive 'single integration strategy' solution. Our flexible data integration approach is a reliable enabler and accelerator for modern data architectures, such as data fabric, data mesh, unified data platforms and hybrid/multi-cloud environments.

Organizations increasingly leverage the Data Virtuality Platform as a solution for delivering an operational foundation for data quality and data governance improvement. Data Virtuality combines data and metadata into one platform, providing a more direct and manageable governance solution.

Key Data Virtuality Platform Features and Roadmap for Supporting and Enhancing Data Governance

The following section highlights some of the current and planned data governance capabilities within the Data Virtuality Platform.

Security and Access Management

Data Virtuality provides secure access based on approved row and column level permission management. All security authorization and updates are fully audited and instantly available for search and compliance reporting.

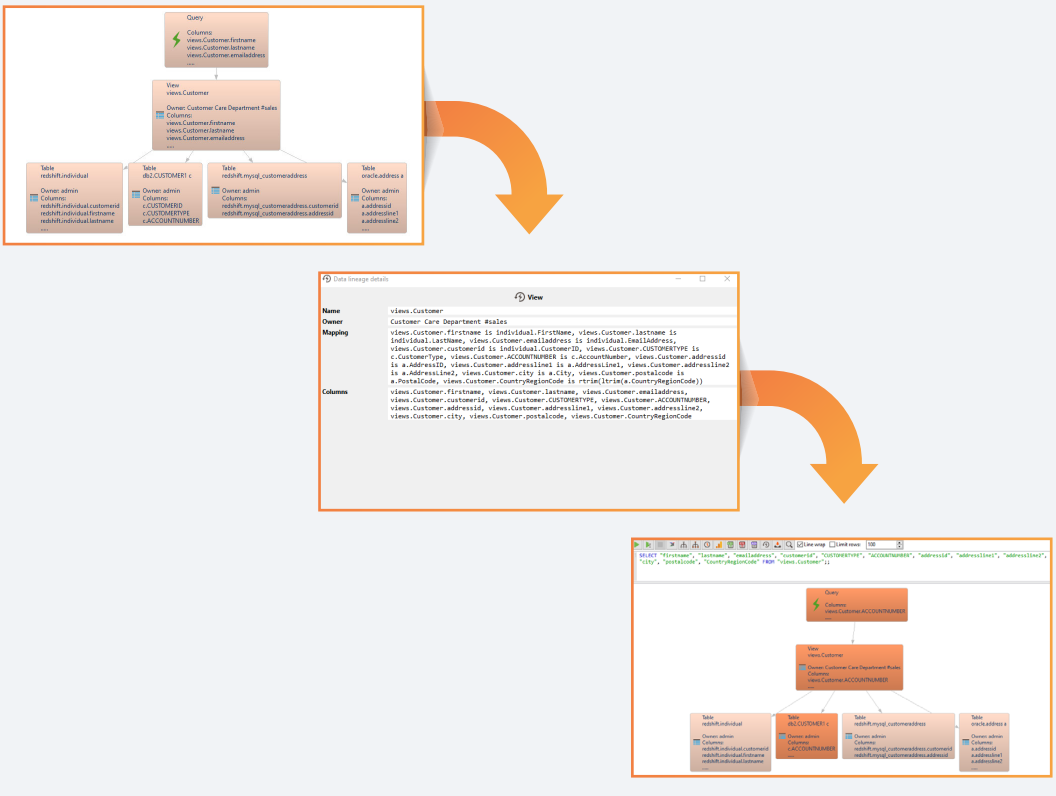
Exchanging Metadata with Data Governance Platforms

The PGGM case study illustrates that the Data Virtuality Platform provides a structured, open and accessible interface with tools such as Collibra and other dedicated data governance products. We have multiple examples of supporting data governance use cases where our data lineage and other metadata have been successfully integrated into commercial data catalog, data lineage and data dictionary solutions such as Collibra and Infogix.

Note: Refer to the PGGM case study below for details of Data Virtuality and Collibra integration.

Data Lineage Management

The Data Virtuality Platform provides instant data lineage, all the way from the top-level dataset view, down to the underlying data flows, and eventually back to the data source. Our approach of visualizing data lineage allows for faster development, testing and accessibility of data services whilst also providing rapid root-cause data defect resolution whenever poor quality data is discovered.



- Data Catalog

  Data Virtuality has its own proprietary data catalog functionality, providing access to simple metadata and descriptions at the view level, as well as customizable attribute-level descriptions and definitions. We are constantly expanding the functionality of our data catalog, so contact us if you wish to learn more about current capabilities and catalog roadmap over the coming months.

- Personal and commercial data masking/obfuscation

  With one of the goals of a mature data analytics and data governance strategy being to 'unlock' a greater volume of data assets for utilization across the wider business, there is a delicate balance between the democratization of data and the need to maintain security and privacy.

  To empower greater accessibility but within a controlled environment, the Data Virtuality platform can obfuscate (hide from view) all forms of sensitive or personal and commercial data, such as:

  - Commercially sensitive trading information
  - Credit card and personal financial data
  - Security and access details
  - Medical information

Case Study: Integrating Data Virtuality and Collibra to Deliver Improved Data Governance at PGGM

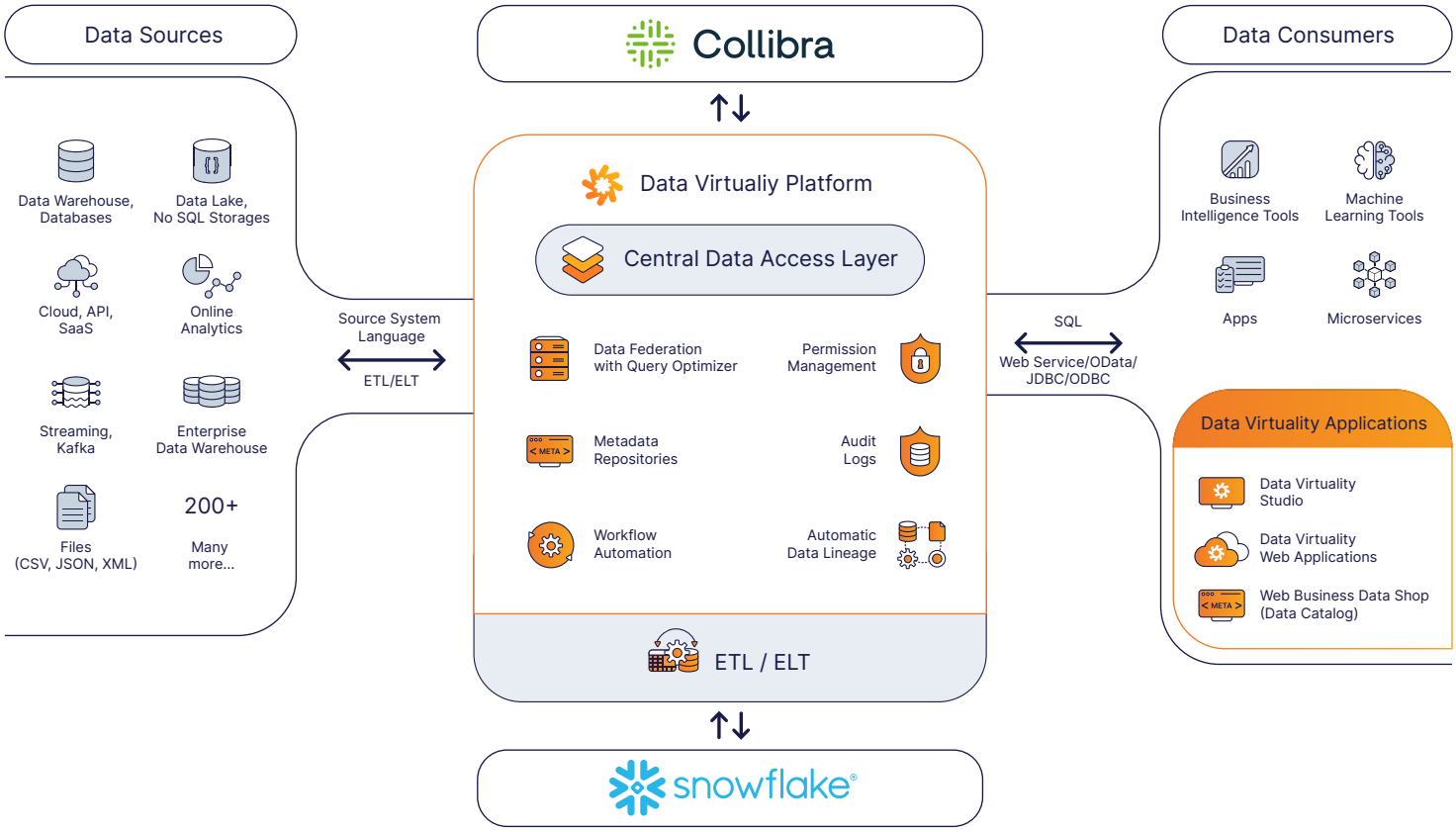
PGGM is one of the top ten pension funds globally, managing nearly €291 billion of pension assets and administered pensions for 4.4 million participants as of December 2021. Before we worked with PGGM, they had experienced some all too familiar data governance challenges:

> We wanted to ensure clear, unambiguous data governance. Before, if you wanted to perform an analysis, you had to know who managed certain data in order to find the information. Moreover, we were not always sure whether the data quality was good and whether users were authorized to use certain data. Due to the lack of clarity ... it could take weeks or even months before requests from employees to use data were approved.
>
> Arjan Surstedt, Enterprise Data Architect, PGGM

To meet the growing demand for advanced analysis and reporting, PGGM selected Data Virtuality, Collibra and Snowflake to help design a modern infrastructure that delivered the required speed, accessibility and governance features expected of a finance business in a regulated sector.

As a result of this combined architecture delivery partnership, PGGM now benefits from:

- A modern data platform in which data is integrated virtually (and physically where necessary)
- A user-friendly data office where employees can easily find, retrieve, and understand the data
- Structured auditing and governance of the definitions, data lineage, and authorization of PGGM data



Next Steps

To see how we approach data governance, the best course of action is to arrange a demonstration of our Data Virtuality Platform. Together with a member of our senior team, we'll discuss your requirements, explore relevant case studies and success stories, then brainstorm creative ways to help you move forward.

Want to learn more? Read the full case study: https://datavirtuality.com/en/customers/pggm/

About Data Virtuality

- Founded:
  2012 by Nick Golovin (PhD) in Leipzig, Germany after 8 years of research
- Offices:
  Munich, San Francisco, Leipzig
- Solutions:
  Data Virtuality Platform SaaS
  Data Virtuality Platform On-Premises
  Data Virtuality Pipes Professional
  Data Virtuality Pipes
- 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

Message: info@datavirtuality.com
Visit: datavirtuality.com
Request Demo: demo@datavirtuality.com
Data Virtuality Platform SaaS Free Trial: https://eu.platform.datavirtuality.com/#/start-trial

data virtuality