denodo

DATA VIRTUALIZATION FOR TODAY'S CDO

Data has always been a valuable asset, but today's chief data officer (CDO) knows that the value of data grows to the extent that the data can be readily combined with other data sets. Today's CDO also knows that with the emergence of new data types, such as data from cloud and big data sources, sensor data from the Internet of things (IoT), and data from social media feeds, as well as data types that were thought to be inaccessible for analytical purposes, such as transactional data and unstructured data, the value of data is potentially limitless.

However, today's CDO is also well aware of the challenges of integrating data, especially between modern and legacy sources, and between data-at-rest and data-in-motion.

Data virtualization is a modern data integration technology that overcomes most, if not all, data integration challenges by leveraging a unique approach: Rather than replicating data, and moving it to a new, consolidated repository, data virtualization connects users and applications to a view of the data, in real time, leaving the source data exactly where it is. In addition to saving on storage expenses, data virtualization enables access to any type of data source, including those that are traditionally not compatible with physical data warehouses.

Gartner has been carefully observing how the data integration landscape is changing. Here, we share key data integration trends that Gartner has published, followed by a brief description of how data virtualization can support that trend, specifically for today's CDO.

"By 2021, 75% of prebuilt reports will be replaced with or augmented by automated insights delivered on a 'most needed' basis." [1]

Data virtualization facilitates the creation of reports on an ad-hoc, or as-needed basis, since data virtualization does not require data to be physically replicated simply to create an integrated report.

Seacoast Bank

Seacoast Banking Corporation of Florida leveraged data virtualization to combine data from a variety of data sources, enabling a true self-service BI infrastructure. The new system enables business stakeholders to generate operational reports and business analysis reports much more quickly than ever before. What would have taken eight months using legacy extract, transform, and load (ETL) processes now takes Seacoast Bank five months using data virtualization, a **40% faster time-to-market**.

“By 2020, most data and analytics use cases will require connecting to distributed data sources, leading enterprises to double their investments in metadata management.” [1]

Data virtualization acts as a universal access layer, connecting users and applications to the myriad data sources underneath. The data virtualization layer itself contains no data, only the metadata necessary for accessing the various sources. As such, it creates seamless access to distributed data sources, and a unified platform for easily managing metadata, to reduce metadata management investments.

INDIANA UNIVERSITY

Indiana University leveraged data virtualization as a key foundation for its Decision Support Initiative (DSI), a wide-reaching information management and delivery project designed to provide timely, relevant, and accurate data to facilitate better decision making within the University. Because the DSI connects stakeholders to distributed data sources, **the university relied on the metadata management capabilities** of the data virtualization layer.

“By 2021, utilization of location-agnostic data semantics based upon new technologies and practices will lower data management and integration costs by 35%.” [1]

By establishing a universal access layer, data vrtualization facilitates the creation of business-friendly semantics that hides access complexities from the user, such as where the data is located or which security credentials are required.

vizient.

Vizient leveraged data virtualization to establish a virtual financial data mart that enables single, seamless data views across a variety of subject areas, unified under business-friendly terms such as “Supplier Sales.” **The data mart successfully serves over 400 active users across 6 departments**.

“Through 2020, organizations adopting data hub strategies will achieve outcomes dependent on shared and governed data with at least 60% lower cost.” [1]

By serving as a universal access point for all enterprise data sources, data virtualization is the easiest, most straightforward, and most cost-effective method for implementing a data hub.

TransAlta

TransAlta leveraged data virtualization to implement a new energy trading system implementation that acted as a data hub, drawing on data from a variety of internal and external sources, some of which were stored in the cloud, and some of which were stored on-site with regulators. The new data hub **reduced unnecessary labor and provided stakeholders with self-service capabilities**, reducing reliance on IT.

“Between 2016 and 2019, spending on real-time analytics will grow three times faster than spending on non-real-time analytics.” [1]

By enabling access to the widest range of data sources, in real time, data virtualization enables real-time analytics.

logitech

Logitech leveraged data virtualization **to integrate cloud and on-premises data sources in real time**, to provide analytics and reporting applications with a single source of truth. The new system enables business users to consume information in an easy, self-service manner and frees them to become “tool agnostic,” since all tools tap into the same unified view of the data. Through rapid prototyping and the optimal usage of resources, the new system has significantly reduced operational expenses as well as costly re-engineering efforts.

[1] Gartner: **100 Data and Analytics Predictions Through 2022**.