Realizing the Promise of Self-Service Analytics

SOLUTION
Self-Service Analytics

WEBSITE
www.denodo.com

PRODUCT OVERVIEW
The Denodo Platform offers the broadest access to structured and unstructured data residing in enterprise, big data, and cloud sources, in both batch and real-time, exceeding the performance needs of data-intensive organizations for both analytical and operational use cases, delivered in a much shorter timeframe than traditional data integration tools.

Data virtualization is empowering business analysts to gain real-time access to integrated data across the enterprise.

Self-service analytics promises to liberate business users to perform analytics without the assistance of IT, and this in turn promises to free IT to focus on enhancing the infrastructure.

Today, numerous desktop analysis tools begin to realize some of the promises of self-serve analytics by enabling business users to slice and dice data and pour it into a variety of full-featured reports and dashboards. However, self-service analytics initiatives are hindered by several key challenges at the data level:

1. **Fragmented Data.** Data is spread across multiple heterogeneous databases, data warehouses, cloud and big data systems, No SQL sources, and flat files.

2. **Multiple, High-Maintenance Data-Integration Initiatives.** When a business user needs to query across multiple heterogeneous sources, companies often charge IT with establishing ad-hoc point-to-point integrations using ETL processes. But if a source needs to be changed, such processes need to be re-written, which is costly and time-consuming.

3. **Data Delays.** It can often take a long time, such as multiple months, to deliver requested data using legacy data integration processes. After such a time lapse, the data is likely to be less accurate and relevant.

4. **Poor Data Integrity.** When business analysts go directly to the data sources themselves, they may not go to the authoritative sources, resulting in data that is of questionable quality.

5. **Untraceable Data Lineage.** Finally, if users collect data from sources directly, they may not keep an accurate record of where the data came from, hindering the ability to determine data quality, and eroding trust in the data.

The reality is that tools alone will not enable true self-service analytics. In fact, no matter the tool, or its features, if any of the above data challenges are present, self-service analytics, by definition, is not possible, since IT will be called upon to deliver a solution. Due to issues of data integrity alone, Gartner says that "Only one in 10 [self service analytics] initiatives will be sufficiently well-governed to avoid data inconsistencies that negatively impact the business." [1]

Data virtualization is a flexible, modern data integration technology that overcomes each of the five challenges above, enabling companies to realize all of the promises of self-service analytics. In this brief, we illustrate how data virtualization enables self-service analytics, and we close with two case studies of companies that have leveraged data virtualization for successful self-service analytics initiatives.

---

[1] "Van der Meulen, Rob. "Managing the Data Chaos of Self-Service Analytics," December 17, 2015 (http://www.gartner.com/smarterwithgartner/managing-the-data-chaos-of-self-service-analytics/)

What is Data Virtualization?

Data virtualization is a data consolidation and integration technology. But whereas most data integration solutions move a copy of the data to a new, consolidated source, data virtualization offers a completely different approach.

Rather than *moving* the data, data virtualization provides a view of the integrated data, leaving the source data exactly where it is. This means that companies do not have to pay the costs of moving and housing the data, and yet they still gain all of the benefits of data integration.

Because data virtualization accommodates existing infrastructure in its existing state, it is relatively easy to implement, compared with other solutions. And because it provides data in real time, from a variety of systems that are normally very time consuming to integrate, such as transactional processing systems and cloud-based storage systems, it can support a wide variety of uses.

Here is how data virtualization overcomes each of the challenges mentioned at the start of this brief:

1. **Fragmented Data is Seamlessly Unified.**

With a data virtualization layer in place, all of the data, across myriad systems in its various formats, appears to users as though it sits in a single, easily accessible repository.

2. **High-Maintenance Data Integrations are Replaced by a Single, Low-Maintenance Data Virtualization Layer.**

Unlike legacy data integration technologies such as ETL scripts, data virtualization can easily accommodate changes to the source data without heavy modification.
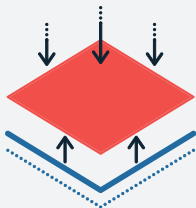
3. **Data Delays are Virtually Eliminated, as Data Can Be Accessed in Real Time.**

Integrated views of the data, even across myriad heterogeneous sources, can be delivered to users in real time.
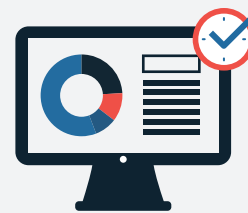
4. **Data Integrity is Preserved.**

Because all of the data sources are accessed through the data virtualization layer, companies can use the data virtualization layer to establish strong governance protocols and specify authoritative sources.

5. **Data Lineage is Fully Traceable.**

Also, because all data flows though the data virtualization layer, data lineage is fully traceable from users to sources.

Case Studies

Denodo's innovative customers Seacoast Bank and Indiana University have successfully leveraged data virtualization to implement true self-service analytics.

Seacoast Bank

Seacoast Banking Corporation of Florida is one of the largest community banks in Florida. Recently, the bank was feeling the impact of maintaining separate systems for such functions as back office operations, data warehousing, and loan origination, and a series of mergers and acquisitions were adding to the complexity.

Seacoast Banking users from Core Banking, Internet Banking, Risk, and other groups had to request static reports from the IT team for operational and analytical purposes. This ad hoc, manual reporting process was both inefficient and time consuming. Seacoast wanted to implement self-service analytics, so business users could freely manipulate the data to meet their needs.

Solution

Seacoast leveraged the Denodo Platform, which uses data virtualization to seamlessly integrate operational data across cloud-based and on-premises information systems and deliver the aggregated views to analytical and reporting tools such as SAS and Tableau. With the data virtualization layer in place, Seacoast is able to provide enterprise-wide self-service BI and develop interactive trend reports using modern reporting tools.

Seacoast is now quickly accessing, unifying, and modeling new data to serve multiple business units, which enables the bank to handle credit administration, risk mitigation, internal operations, and Bank Secrecy Act requirements, in near-real-time. Seacoast analysts can now quickly pull data from across the enterprise, which the bank regards as a significant improvement.

Mark Blanchette, VP and director of Business Technology and Data Management at Seacoast Bank, says that "Denodo's data virtualization technology has played the most important role in enabling our business users to garner valuable information through self-service reporting. The Denodo Platform's capabilities have significantly increased the speed at which business is carried out at Seacoast Bank."
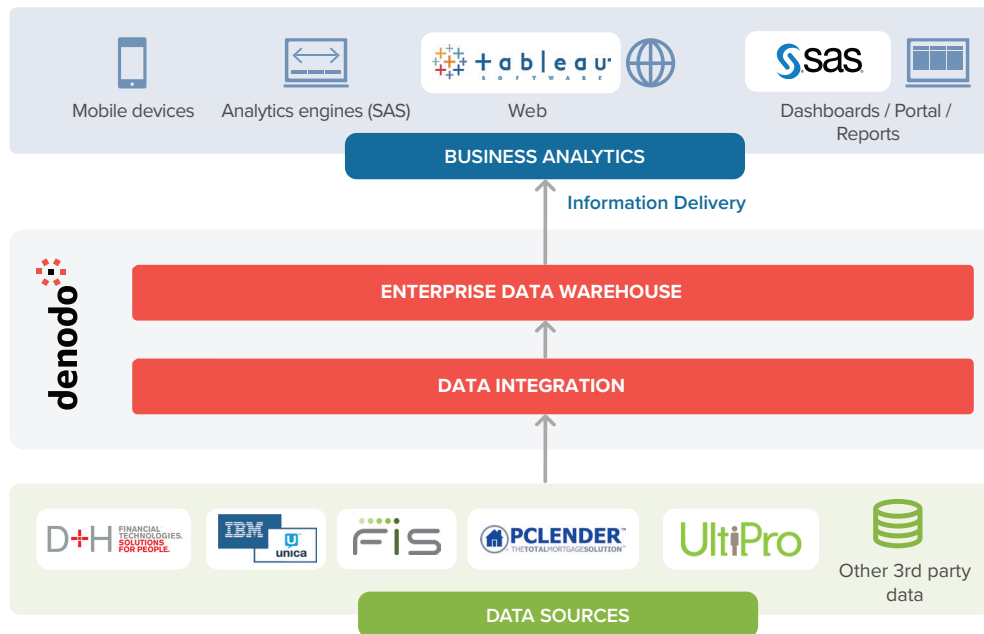
Benefits

- The data virtualization layer has significantly improved the ability of Seacoast's Banking operations groups, such as Deposit and Loan Operations, to make timely, accurate decisions.
- Data virtualization integrated the data in less than half the time that a traditional ETL solution would take, enabling Seacoast to meet the operational and analytical needs of multiple business units within the organization.
- Data virtualization helped Seacoast to reduce reporting time from up to three days for static reports to as little as two hours for interactive, self-service reports that serve business users in loans, deposits, fraud, credit, and risk departments.
- Powered by data virtualization, Seacoast can now perform critical business operations, such as loan processing, in real time.

Indiana University

Indiana University (IU), with eight campuses, serves more than 114,000 students with more than 19,000 employees. The institution has had a long history in analytics, dating back to the days of the mainframe, focusing on such factors as hiring trends, employee performance, class sizes, and student success. Recently, IU began an analytics modernization project called the Decision Support Initiative (DSI), dedicated to increasing the availability of timely, relevant, and accurate information throughout the university.

Historically, data and its corresponding business logic were stored across multiple, siloed systems, making it extremely time consuming to gather and combine the relevant information decision makers needed. In some cases, data activities failed entirely because stakeholders could not find the data elements they needed, or common definitions of sources-of-record, which IU did not keep. The university's data integration toolset, primarily built around ETL processing, required broad skillsets to deploy, maintain, and manage, and resources with these skills were scarce. As a result, the university was dedicating too much time to information-access development; by the time data was retrieved, it was often less useful, or even irrelevant, for decision-making.

Solution

IU leveraged the Denodo Platform to create a single, unified system of easily consumable information. The Denodo Platform uses data virtualization to seamlessly connect the university's diverse sources, including Oracle, MS-SQL, Amazon Redshift, web services and Box.com, and provide it in real time to Tableau, Excel, and other analytics applications. By establishing a layer for unifying data from the myriad sources, the Denodo Platform also enables the implementation of security and governance protocols across the entire infrastructure from a single point of control, which facilitated compliance initiatives.

Users can now access and manipulate the data from a Web front end, in a seamless, self-service manner, and IU decision makers now have ready access to the data they need, when they need it.
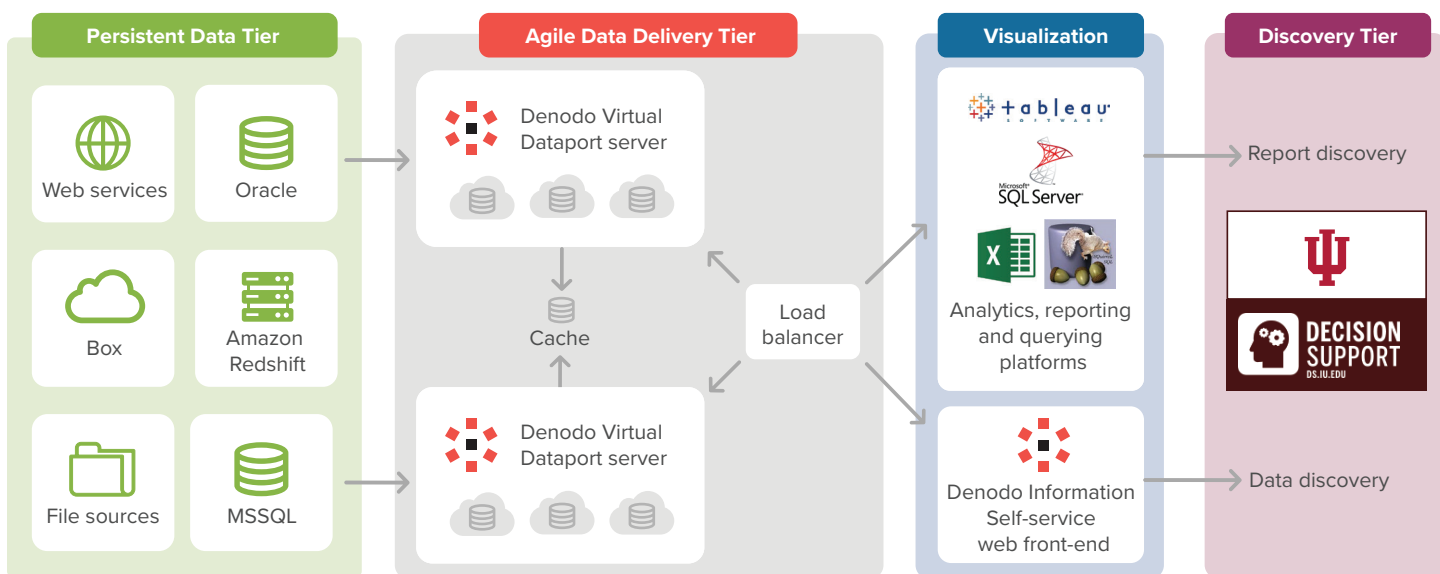
Benefits

- Data virtualization has significantly improved information agility across the university. Data can now be defined and accessed almost instantaneously, no matter where it resides and with minimal effort.
- Diverse data spread across the enterprise can now be accessed securely with a proper authorization structure. Rules can be applied no matter when the data is accessed or where the data is stored.
- Core business intelligence logic is being centralized, reducing the duplication of efforts and enhancing development efficiency.
- Indiana University now has a searchable data dictionary, which helps report writers find the data they need and help improve the self-service experience.