

Creating a Single Source of Truth



Overview

After reading this [McKinsey article](#) detailing how three B2B companies successfully increased their sales, we felt inspired to delve deeper into one of their transformation strategies: creating a single source of truth (SSOT) by centralizing data into a warehouse. The article painted a picture of this approach as a key driver of success, but in reality, the journey to a true SSOT is far more nuanced.

In an era where data-driven decisions are not just an advantage but a necessity, achieving a single source of truth (SSOT) has become the holy grail for businesses. While centralizing data into a warehouse seems like a straightforward path to SSOT, the reality is much more complex. Let's explore why mere centralization isn't enough and the strategies companies can adopt to ensure a truly effective SSOT.

Top 5 Reasons Why Centralization Alone Doesn't Cut It

- Data Diversity
- Data Quality
- Data Integration
- Dynamic Data
- Data Governance

Data Diversity

Data diversity presents a significant hurdle as information comes in various formats, from structured databases to unstructured text, each requiring different handling. Simply dumping all this into a central warehouse doesn't magically align these diverse data types into a harmonious dataset.

Data Quality

The issue of data quality cannot be overstated as data discrepancies, inaccuracies, and duplicates can plague centralized repositories just as they effect decentralized ones. Without addressing these quality issues, businesses risk making decisions based on flawed data.

Data Integration

Integration challenges further complicate matters as seamlessly merging data from disparate systems and sources requires sophisticated integration tools and strategies. This goes well beyond the act of centralization.

Dynamic Data

Businesses are not static, and they continuously evolve, their processes change, and new data sources emerge. What was once a centralized solution can quickly become outdated or insufficient.

Data Governance

Data governance emerges as a critical challenge in the process of centralizing data, aiming for a single source of truth (SSOT). While centralization attempts to bring data together, ensuring consistent and effective governance across this unified dataset can be complex and very challenging.

Data Products as a Reliable Approach to Achieving SSOT

Building data products is an innovative approach to solving the challenges associated with achieving a single source of truth (SSOT) and improving data governance. Data products go beyond traditional data management practices by encapsulating data in a way that is directly usable, valuable, and actionable for end-users, whether they are internal stakeholders or external customers and partners. This approach significantly enhances how organizations manage and leverage their data for decision-making, innovation, and competitive advantage.

What are Data Products?

Data products are operationalized datasets that are packaged and designed to be reused, shared, and applied to solve specific problems or deliver insights. They treat data as a product, focusing on the user experience, data quality, and accessibility. Data products can service internal analytics dashboards and machine learning models to customer-facing applications that use data to provide personalized services or recommendations.

How Data Products Address SSOT Challenges

- Detailed Data Governance
- Decentralized yet Consistent Data Management
- User-centric Approach
- Facilitates Scalability and Innovation
- Enhances Data Literacy
- Enables Reusability

Detailed Data Governance

Data Governance should cover all aspects of access policies, data masking and apply that dynamically without having to move data from one system to another. The data governance should integrate with active directory or equivalent systems.

Decentralized yet Consistent Data Management

Data products can support both centralized and decentralized ownership approach. In a decentralized approach to data management, different teams within an organization can develop and manage their own data products while in a centralized approach, a central data team manages the data product life cycle. However,

data products should be built with a common set of standards and frameworks to ensure consistency and integration of data across the organization.

User-centric Approach

Focusing on the end-user experience encourages the development of intuitive interfaces and access patterns for data which helps democratize data access and usage across the organization. This user-centric approach ensures that data products are designed with the specific needs and contexts of their users in mind, promoting broader engagement with data and insights.

Facilitates Scalability and Innovation

As organizational needs evolve; data products can be iteratively improved and scaled to meet new requirements. This flexibility supports innovation, as teams can rapidly prototype, test, and deploy new features or products based on real-time data insights.

Enhances Data Literacy

The development and use of data products encourages a culture of data literacy within the organization. As users interact with data through well-designed products, they become more comfortable and skilled in leveraging data for decision-making.

Enable Reusability

Data products revolutionize the reusability of data assets within an organization. Traditionally, data resided in silos, making it difficult to combine and utilize across different projects. Data products break down these silos by creating pre-built, modular components that encapsulate specific data sources, transformations, and functionalities. These components can be easily integrated and reused in different workflows and applications.

This reusability fosters efficiency reduces redundancy in data preparation efforts and ensures consistency across various data-driven initiatives. Imagine a data product providing customer segmentation insights – this same product, with minimal adjustments, can be leveraged by marketing teams for targeted campaigns and by sales teams for personalized outreach. Data products empower a "build once, use many times" approach, maximizing its value.

Implementing Data Products

To successfully implement data products, organizations should:

1. Define Clear Objectives
2. Establish Cross-functional Teams
3. Adopt Agile Methodologies
4. Invest in a Data Product Technology Platform

1. Define Clear Objectives

Identify specific business needs or user problems that data products can address. This ensures that the development of data products is aligned with strategic goals.

2. Establish Cross-functional Teams: Data Product

development requires collaboration across data scientists, engineers, product managers, and business stakeholders to ensure that the product meets technical standards and business needs.

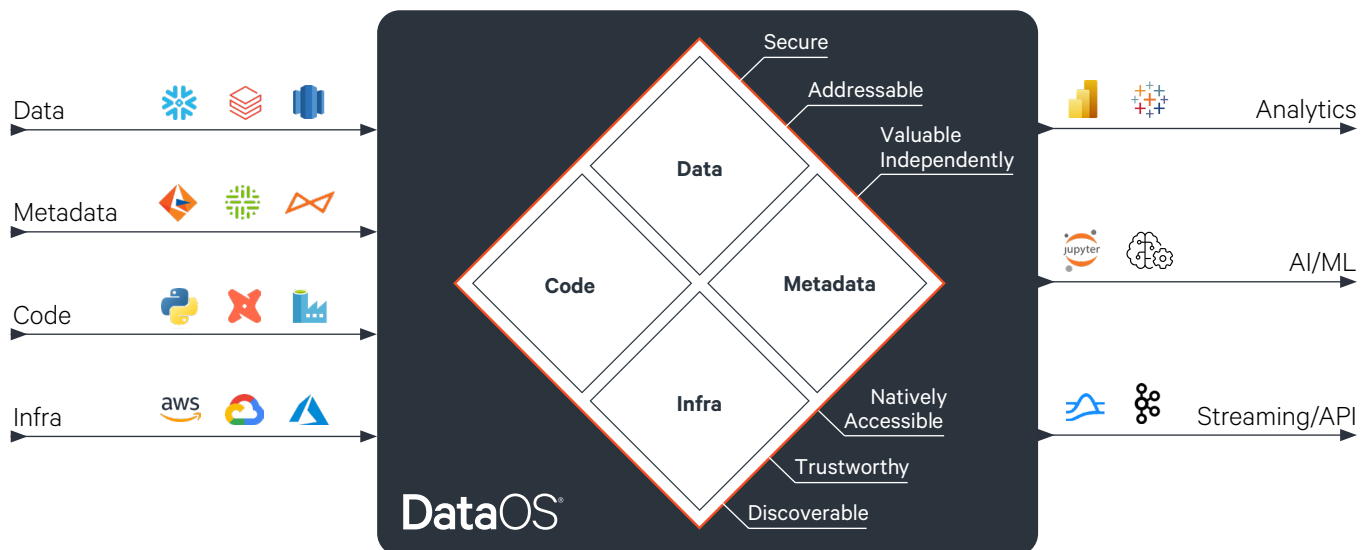
3. Adopt Agile Methodologies

Use agile development practices to iteratively build and refine data products based on user feedback and changing requirements.

4. Invest in Data Product Technology Platform

It is very likely that your current data infrastructure was not designed to facilitate creation of data products and will need a lot of skills and resources to build data products from scratch. So, it may be the right time to evaluate a platform like [DataOS](#) that is designed to develop data products on top of your existing infrastructure.

Data Product should be created like a docker container so that it can run in any cloud. This means it should include not just the data but also the metadata, transformation logic (code) as well as the infrastructure to run it.



What Makes a Good Data Product?

Here are the eight characteristics of a good data product in no particular order:

Discoverable

Available in a centralized catalog with an easy way to search

Understandable

Documented with clear purpose, schema, metadata, business context and usage guidelines

Addressable

Secure and scalable access mechanisms such as APIs and web interfaces

Secure

Enforce security and privacy protocols through fine-grained access controls

Interoperable

Should interoperate with existing data (Snowflake, Databricks, SQL Server), metadata (Alation, Unity), transformation code (dbt, Python, Azure data factory) or infrastructure (AWS, Azure, GCP)

Trustworthy

Establish trust through service level objectives (SLO)

Natively Accessible

Directly accessible through a self-serve platform, allowing consumers to easily explore, analyze and visualize data products.

Valuable on its own

Autonomous entity encapsulating all necessary components including code, infrastructure, policies, and documentation needed.

Solve SSOT Challenges with Data Products

Building data products represents a strategic approach to solving the complexities of achieving a single source of truth and enhancing data governance. By focusing on creating valuable, user-centric data solutions, organizations can not only improve their data management practices but also unlock new opportunities for innovation and growth.

[Visit Modern Resources](#) →

[Schedule a Demo](#) →



The Modern Data Company

The Modern Data Company, headquartered in Palo Alto, California, empowers organizations to unlock the true potential of their data with DataOS®, the world's first fully integrated, extensible data operating system to build data products. DataOS® simplifies data management by unifying silos, bridging the gap between conceptual and physical modeling, and offering ontology support, graph modeling, and a virtual data tier. This unique approach streamlines data product creation and deployment, enabling businesses to avoid lengthy technical implementations and drive faster iterations of data-driven strategies. By eliminating the need for complex data infrastructure and facilitating seamless integration with existing systems, DataOS® accelerates time-to-value and frees businesses to focus on achieving their strategic goals. Learn more about how The Modern Data Company can help your organization harness the power of data at themoderndatacompany.com/dataos.

