



Guide to

Data as a Product in Snowflake



Passerelle



How to Support Data as Product in Your Organization

It's time to start thinking about your Data as a Product.

The data being created in your organization is only growing, along with the tools that both manage and create more data. It's an alphabet soup of data sources – CRMs, ERPs, POSs, HRMSs, LMSs, R&D, FP&A – and inside each acronym lies a treasure trove of information on your operations, your customers, your sales and marketing, and your human resources.

By thinking about your Data as a Product, organizations make the shift from considering data as a tool or a byproduct to treating it as a valuable strategic asset.

What is Data as a Product? What is a Data Product?

So what is **Data as a Product**, and how is it different from a **Data Product**?

Imagine going to your favorite bookstore – but instead of books, the shelves are stocked with ready-to-use datasets. Just as you wouldn't find a bookstore full of unfinished manuscripts, datasets in your data store are created after data is collected, processed, refined, and transformed into a form that is usable





and valuable to its end users. As the amount of data grows in your organization, the shift in mindset that enables Data as a Product becomes imperative. It's not enough to collect and store your data – organizations must ensure data is accurate, accessible, and fit for its purpose. By supporting Data as a Product, your organization recognizes the intrinsic value of data as a strategic corporate asset. In this analogy, **Data as a Product** is the methodology that supports self-service, discoverable and usable assets. The **Data Product** is the consumable asset itself, and a by-product of the Data as a Product approach.

Laying the Foundation for Data as a Product

How you store your data will determine whether you can support Data as a Product methodology. In a traditional data warehouse, transformed data is stored in a single, structured format that is optimized for read-heavy operations. While this traditional data warehouse architecture can be trusted to consolidate data and support business intelligence initiatives, it comes with an IT dependency that limits self-service insights and data discoverability. An approach such as a data lakehouse combines the benefits of warehouse storage with features associated with a data lake – adding a storage layer that can handle unstructured and semi-structured data. This makes it easier for users to create, discover, and use data products independently and supports the creation of Data Products.





WHY DATA AS A PRODUCT?



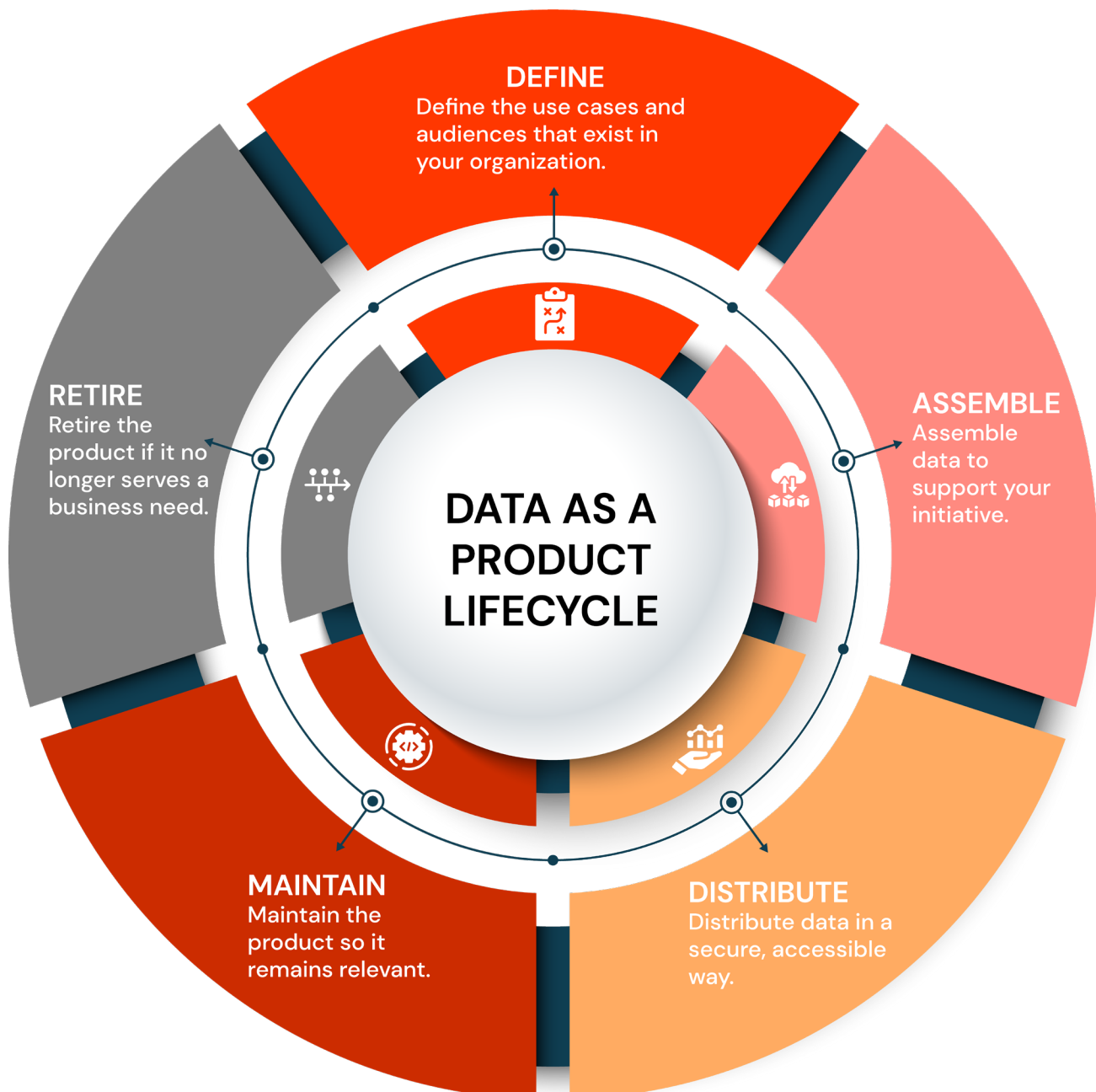
By viewing your Data as a Product, organizations can develop a framework for mining the value of data – defining what use cases and audiences exist in your organization, assembling the data to support your initiative, distributing data in a secure, usable and accessible way, maintaining the product so it remains relevant and retiring the product if it no longer serves a business need. When your organization manufactures a Data Product, you give Data Consumers the opportunity to know and rely on data in their day-to-day operations. A Data Product creates a tacit contract between the Data Consumer and the Product Manufacturer – in essence, creating the same kind of trust, brand loyalty and customer satisfaction feedback loops that every organization strives to achieve in product and service delivery.

Data Products are only one outcome of embracing a Data as a Product mindset. A Data as a Product approach yields other significant benefits, including improved data quality and enhanced data governance, as organizations implement high standards for trustable data. This approach fosters governed data with increased accessibility and data democratization as data is moved into self-service, user-friendly formats. Thinking of Data as a Product creates new revenue streams through data monetization, and provides a competitive advantage by enabling the development of new products, services, and business models.

Your Data as a Product Toolkit

In this article, we'll outline a modern, cloud-native tech stack that can simplify and automate your process your Data as a Product approach. These tools create replicable, scalable systems with best-practice data management.

- **Define:** ALTR, Qlik Talend, Snowflake and Data Rocket®
- **Assemble:** Qlik Talend, Snowflake, and Data Rocket
- **Distribute:** Snowflake, ALTR and GrowthLoop
- **Maintain:** Qlik Talend, Snowflake, ALTR and Data Rocket
- **Retire:** Snowflake and ALTR





Define: Create a Roadmap of Data Product(s)

Let's start with the eponymous principle that guides a Data as a Product Approach: Organizations must treat their data just as they would any product they want to develop, sell, or maintain. Just as it wouldn't create a new service offering, widget or incentive program without significant R&D and planning, an organization can't embark on a new data initiative without clearly articulating the final vision – in this case – a Data Product.

To create a roadmap for a Data Product, organizations need to consider how the data is used, who will use the data, how data will be collected, how the Product will be distributed and consumed, and how it will be maintained and retired.

As an example, here are some common Data Products and associated benefits to an organization.

- **Customer 360-Degree Insight:** An aggregation of customer attributes and operational behavioral data from various sources to provide comprehensive analytical insights into each customer. This Data Product helps organizations better understand customers to enhance personalization and customer engagement and improve customer service across all interactions. With Customer-360 Insights, organizations can get a predictive understanding of the

potential risk of customer churn and mitigate the risk of losing a customer with data-informed prescriptive data insights.

- **Sentiment Analysis and Customer Feedback:** Analyze customer reviews, social media posts, and feedback to gauge sentiment and identify trends. Sentiment analysis can help gauge customer opinions and improve products and services based on feedback.
- **Predictive Analytics for Sales and Demand Forecasting:** Create a predictive analytics solution that forecasts sales and demand using historical data and advanced algorithms. This Data Product supports strategic planning and decision-making by predicting future trends and demand patterns.

When you know what Data Product you would like to build, organizations can create the foundation for production by classifying your data and defining the technical requirements.

Classification – ALTR

To get started, you must understand exactly what data you possess, where the data lives, and which data is sensitive. As data continues to flow into the business, you need to make sure that sensitive data is well governed. You don't have time to check your data row by row – that's practically impossible. Yet, knowing what data is sensitive is the foundation for healthy data.

With ALTR, you can facilitate data classification with Snowflake Classification or Object Tags or Google DLP and receive results in minutes. If sensitive data isn't identified, it's impossible to protect, leaving gaps in both privacy and security. ALTR integrates data classification into the policy enforcement engine, allowing YOU users to automatically find, tag and enforce governance policy on data easily, all from the ALTR interface.

Define and Design – Qlik and Snowflake

With classified data identified, you can define and design the Data Product – ultimately determining how data is structured, stored and accessed in Snowflake.

To achieve this, start with a data model design and maintain the data architecture across your data team. Once the data modeling phase

is complete, integrating and operationalizing the data becomes the next critical step. With Qlik Talend Data Management, organizations can extract data from various sources, transform it as needed, and load it into Snowflake. Talend's robust data integration capabilities streamline this ETL process, supporting complex transformations and ensuring data quality and consistency. This integration helps maintain the fidelity of the data model throughout the data lifecycle and enhances the performance and scalability of the Data Product.



Assemble: Collect Data to Support the Product

Production – Qlik, Snowflake, Data Rocket

Once an organization has defined the desired Data Product design, it must collect data to ensure it is cleansed, integrated, transformed, validated, and trusted. This process is crucial in establishing a reliable and robust Data Product. To facilitate this complex data management challenge, Data Rocket accelerates and optimizes the ingestion, integration and transformation of transactional data to an analytics-ready state into Snowflake.

Using Qlik Talend Data Management, Data Rocket automates the ingestion of any data from any data source. Data Rocket's Observe and Control dashboard provides historic and real-time data ingestion processing information to help troubleshoot ingestion problems at the source.

Once data is ingested, Data Rocket integrates disparate data sources to create a unified dataset that is ready for transformation and analysis. To support data quality and creation of a single version of the truth, Data Rocket comes with Data Quality Watch® – a data profiling tool that measures data on Completeness, Accuracy, Consistency, Validity, Timeliness, Popularity and Integrity that lets organizations can create their own Quality measure specific to their business. A Mastered Data Framework automates

ASSEMBLE AND DISTRIBUTE



a simplified data mastering process.

With trusted data in place, Data Rocket's metadata framework supports the transformation of data into the data marts that will ultimately populate your Data Product, with transformation models that fit your organizational needs.

Distribute: Make Sure Data is Secure, Usable and Accessible

Distribution – Snowflake and ALTR

The Data Product needs to be accessible and secure, available to the right users while safeguarding sensitive information. ALTR, a SaaS platform that sits directly on the Snowflake Data Cloud, can help organizations distribute a data product with confidence by making it easy to set access control policies and usage limits quickly and effectively.

ALTR's unique combination of features – secure data tokenization, user access controls, and real-time alerting – sends notifications if classified data is accessed and limits data access by quantity – even to admins.

Consumption – Snowflake, GrowthLoop

The Data Product should be user-friendly and presented in formats that are easy for end-users to understand and utilize. Channels should be established for users to provide feedback on the Data Product, allowing for iterative improvements.

For marketing Data Products, GrowthLoop is a Snowflake Native App that sits on top of an organization's Snowflake Data Cloud. The platform provides a simple, self-serve interface that gives business users access to all data needed from feedback, sales orders and service to create tailored and personalized messaging to its vendors. Within the same workflow, users can segment highly-targeted audiences, orchestrate personalized cross-channel customer journeys that ensure customers receive the right message at the right time, and measure results to optimize for the future. Personalized messaging that offers the additional product or service the vendor needs will drive significant upsell revenue opportunities, as well as customer loyalty. Better yet, by fully leveraging your Data Product, you'll be able to reach every audience with the robust messaging that best suits their unique needs.



Maintain – Ensure Data Product Remains Relevant

Maintenance – Qlik, Snowflake, ALTR, Data Rocket Observability



Different versions should be tracked and versioned as the Data Product evolves, ensuring users can access historical data or understand changes over time. Updates and patches should be performed to ensure the Data Product captures data updates, transformation, and error corrections.

ALTR monitors and records each request for governed data, along with user, time, and values returned, allowing organizations unparalleled insight into what is happening with their data through data usage heatmaps and analytics. With ALTR, administrators can quickly obtain detailed visibility into data use over time so they can understand patterns, spot abnormalities, and adjust policy quickly.

Retire: Know When Product No Longer Supports a Business Need

Retirement – Snowflake and ALTR

When a Data Product reaches the end of its lifecycle, it should be phased out systematically, with adequate notifications to users and stakeholders. By archiving the Data Product, it can be used for historical reference but doesn't clutter your data ecosystem. The same goes for new version releases.

Snowflake lets you set data retention policies that can be customized based on regulatory and business requirements and provide a cost-effective way to store archived. With Snowflake, you can set retention periods for different types of data so you know data is only kept as long as needed and automatically deleted when it is no longer required. ALTR builds on Snowflake's capabilities with automated access controls that help make sure data is protected, even in retirement.

Data as a Product – Not If, but When

Embracing Data as a Product approach will only become more essential as organizations move to advanced applications – including analytics, ML and AI applications. For most organizations that want to use data strategically, the move to Data as a Product isn't if, but when.

Need help getting started?

Contact us for a complimentary consultation – use the QR code to or email: info@passerelledata.com.



About the Partners in this Guide

Passerelle

Passerelle connects data to action. Our purpose-driven engineering supports Agile Data Governance and AI-Readiness. Our leading-edge technology partner eco-system covers the entire data value chain, spurring adoption through our IP and blueprints, technical expertise, and use-case-based deployment.

In addition to engineering and system integration services, Passerelle is the creator of Data Rocket®, an end-to-end acceleration architecture that modernizes data infrastructure and delivers critical business insights – securely and accessibly. Data Rocket unlocks industry-best data technology for businesses of any size, with a focus on data quality, scalability and advanced data applications. Learn more at passerelledata.com.

Snowflake

Snowflake enables every organization to mobilize their data with Snowflake's Data Cloud. Customers use the Data Cloud to unite siloed data, discover and securely share data, power data applications, and execute diverse AI/ML and analytic workloads. Wherever data or users live, Snowflake delivers a single data experience that spans multiple clouds and geographies. Thousands of customers across many industries, including 647 of the 2023 Forbes Global 2000 (G2K) as of October 31, 2023, use Snowflake Data Cloud to power their businesses. Learn more at snowflake.com.

GrowthLoop

GrowthLoop is the Composable Customer Data Platform (CDP) that empowers marketers to efficiently build audience segments, orchestrate cross-channel journeys, and assess campaign results directly through the Data Cloud. In contrast to traditional MarTech Stacks, characterized by closed ecosystems and data redundancy, the New MarTech Stack is open, orchestrated, and committed to zero data duplication. GrowthLoop serves as the gateway to this innovative landscape, offering centralized control over data and intelligence, freeing teams from vendor lock-in, and fostering an environment of channel innovation. Learn more at growthloop.com.

ALTR

ALTR is the only Data Access Governance and Data Security Platform available as a cloud-native service in Snowflake. Delivered as a pure SaaS solution, ALTR offers near-immediate time-to-value without the heavy professional services lift required by competitive solutions. ALTR's point-and-click user interface greatly reduces the complexity of leveraging Snowflake's powerful native governance capabilities like dynamic masking, row-level security and tag-based policy, all of which can be automated at scale using ALTR's Management API. Through the architectural advantages of its cloud-native integration with Snowflake, ALTR can provide real-time capabilities unavailable in Snowflake itself, like real-time data access observability and real-time access controls. Further, ALTR's alert and notification framework allows InfoSec teams to get real-time signals for out-of-policy requests for data, policy anomalies, and data access observability. ALTR helps Snowflake customers greatly accelerate their time to value through reduced complexity, increased efficiency, and more effective risk management. Learn more at altr.com.

Qlik

Qlik is a global leader in data integration and data management. Qlik Talend Data Management offers a complete, end-to-end solution for modern data management with extensive functionality for data integration, data quality, and data governance. With Talend, organizations have the most flexible solution for working across any cloud, with any data technology, and within any data environment. Over 7,250 customers trust Talend to drive real value for their business. Learn more at qlik.com.



Copyright Passerelle Corp
2024
All Rights Reserved