



# Unlock the Power of Data as a Product

How an end-to-end data strategy supports productization



# Table of contents

What are data products and what can they do for your business? .....	4
Who drives data productization? .....	8
What are the elements of curation? .....	10
What are the elements of an end-to-end data strategy? .....	11
Unlock the value of your data with AWS and AWS Partners .....	13
Conclusion: From data sets to data products.....	14

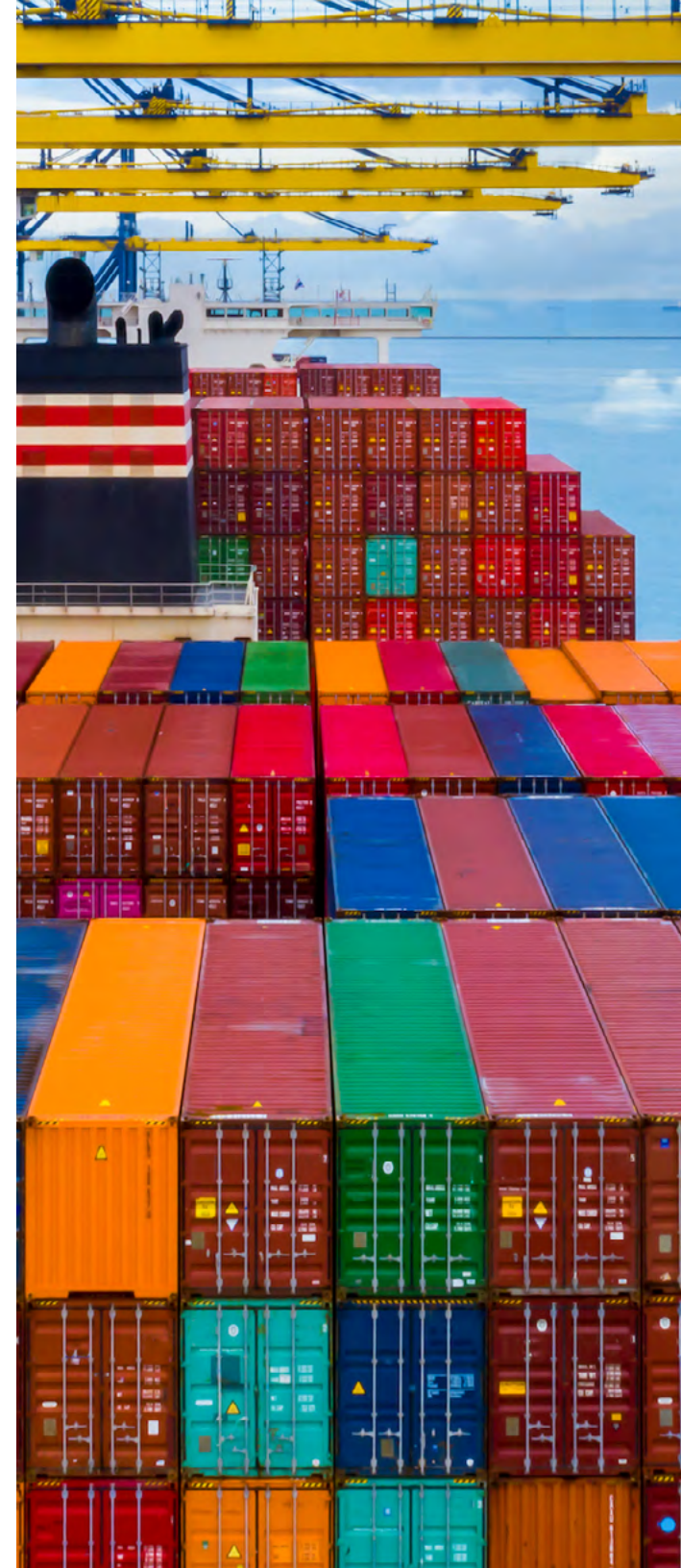
# Introduction

How is data stored and accessed in your organization? How much do you trust your data when it's time to make a decision?

The data landscape for many enterprises is riddled with silos and often characterized by uneven and inadequate access, governance, and quality. To become a data-driven enterprise, you need to remove those silos and open the door to greater understanding and trust in your data.

This is where data productization comes in. When data is treated as a product, you can access all the information you need to make the right decisions. You no longer have to guess what your customers want. Your data resides in an open and standardized format in the cloud. When it's time for analytics, you can be confident that you have a trusted source from which you can build reports, dashboards, visualizations, and predictive models.

This ebook explains the concept of data as a product, highlights an end-to-end data strategy that supports data productization, and describes how AWS Partner solutions can address gaps in your journey to becoming a data-driven enterprise.



In a recent TDWI (Transforming Data With Intelligence) survey, data professionals ranked the top causes for problematic decisions as follows:

**72%**

said rogue data is the culprit

**56%**

named self-service data practices

**33%**

listed data quality issues associated with analytics sandboxes

**28%**

cited data lakes<sup>1</sup>



## What are data products and what can they do for your business?

By making just 10 percent more data accessible, a typical Fortune 1000 company will see a \$65 million increase in net income and a 48 percent reduction in their total cost of operations.<sup>2</sup> So how do you make that happen?

The answer is by improving data productization. When data is perceived as a product, it changes the fundamentals of how an organization contemplates data. While “data as a product” terminology is commonly associated with a data mesh architectural pattern, data products themselves are found across the spectrum of data and analytics. Improving and enhancing data product structures, access, and policies can help all organizations realize near-term and long-term data value.

**So, what is a data product exactly? And what outcomes can data productization drive for your business?**

<sup>1</sup> James Kobiellus, “The Role of Modern Data Governance in an Internal Data Marketplace,” TDWI, 2021.

<sup>2</sup> Brent Dykes, “The Four Key Pillars to Fostering a Data-Driven Culture,” Forbes, 28 March 2019; Ritu Jyoti and Harsh Singh, “The Business Value of AWS Data Lakes, Analytics, and ML Services,” IDC, 2020.

# Data product defined

A data product is any self-contained data entity used to solve a business problem or that can be monetized. It needs to be complete, valuable, high-quality, and ready to deliver precise and accurate insights and analytics. Data products move organizations away from siloed data and take the use of organizational knowledge and “gut feelings” out of the decision-making equation.

Data products fall into three main categories. One is data-derived insights, which are data-product-led conclusions from analytics, dashboards, and visualizations. A second is data-driven services. These are data products in an end state, either as monetizable assets or consumed as a core part of a service, such as transactional data streams. The other category is data-driven enhancements. They use machine learning (ML) to improve and augment analytical models that are components of a larger solution.



## Examples of data products

Here is an example of a data product in each category:

### Data-derived insights

Contextually aware advertising derived from applied ML to user behavior data

### Data-driven services

Visualized sales trends from customer relationship management (CRM) system data

### Data-driven enhancements

Fault detection on an assembly line powered by machine-learning-driven anomaly detection based on manufacturing process data

# The key business outcomes that data products deliver

Data products can help your business in many ways. Here are six key benefits.



## Make better, faster decisions

Data products remove data silos, improve data quality, and address skills shortages. They empower employees, regardless of job function or business unit, to analyze and use data to make decisions.



## Improve customer experience and loyalty

Data products help your business get a 360-degree view of your customers. You can analyze that view to understand them better and improve their experiences. You can even create monetization opportunities with personalized products and services, increasing your chance of retention.



## Optimize business processes and reduce cost

Data products deliver what you need to learn from the past, monitor the present, and prepare for the future. You have the right data so you can forecast effectively, optimize, and automate processes such as the supply chain, and reduce operational costs.



## Handle customer data with care and integrity

The rigorous governance delivered by data productization helps you handle sensitive customer data with care and integrity while managing compliance with privacy regulations and mandates.



## Stay ahead in the market

With data products, your business can continuously refine ML models and business processes. You can then deliver innovative, market-leading solutions based on quantified insights.



## Prepare for the future

A data product is an organizational store of value. Ensuring the right resources have access to data that they trust and that is consistent with standards drives improved organizational alignment. You gain a faster on-ramp to innovation.

**How do you drive data productization in your organization to realize benefits?**  
**Start by working backwards. Define who will use and who will drive your data products.**

# Who uses data as a product and who drives productization?

Applications and people can both process and use data products. *Data producers* collect, process, and store data from their data domain, in addition to monitoring and ensuring the quality of their data assets. *Data consumers* use the data from producers.

Application and application owners, analysts, data scientists, business owners, and operational, security, and development resources aren't the only ones who use and benefit from improved standards and practices for data products. Almost anyone in your organization—from knowledge workers to decision-makers—engage actively or passively with data regularly. In some cases, it's constantly. These resources must have the right product to take responsibility for the transformation, accessibility, processing, and availability of the data. Therefore, you need to ensure you have the correct policy and organizational alignment.

When you are planning your strategy for data productization, you need to focus on what these data users want and how they work. This requires collaborative discovery work, where users tell you the key business questions they need to answer and the outcomes they are seeking. For example, a marketing team might want to use data to optimize account-based-marketing campaigns to increase sales-qualified opportunities. An IT team might want to use data from threat analytics to pinpoint anomalies and better protect their cloud workloads.

How users work is also important. As part of discovery, uncover their preferred methods of consuming data. If it's manual, such as spreadsheets, think of ways to deliver the data in that familiar format but without the silos. Rely on documentation and use definitions and language that people understand and use every day.





## CIOs, CDOs, and other executives

Executives have the power to influence culture, finances, resources, and the acceptance of new data practices. To get their buy-in, you can show how data productization aligns with your organization's shared goals and strategy.

To start, consider what you want data products to do for your business. Much like your work with data users, this includes a period of discovery to determine success and how to measure it. Perhaps you want to achieve cost savings, increase productivity, or foster innovation. Then, establish key performance indicators (KPIs) that show the business value of data productization, a total cost of ownership for data products, and an organizational return on investment. You can also develop a proof-of-concept, set measurable KPIs, and track progress. Collect the metrics and use them to make your case to executives.

## Data stewards

Data stewards are responsible for the overall data vision for an organization or a line of business. They own and maintain databases, data processes, the data roadmap, and data priorities. Their focus is on implementing and ensuring data governance and access controls, along with the process of aligning data to business requirements. They play a big role in developing an end-to-end data strategy for data productization.

## Data curators

Data curators are key to the success of data productization. They make sure that data and metadata are complete, well-described, and in the format and structure needed for easy long-term access, discovery, and reuse. The goal is to deliver more context to users by enabling them to find the right data relevant to their role, which is critical to data-driven decision-making. Data curators understand where and how data is generated, as well as where it's stored. They also ensure that people recognize how to use that data when they find it.

Data curation includes elements that help drive the success of data productization.

# What are the elements of data curation?

Data curation organizes and integrates relevant metadata into a data catalog. This catalog enables data consumers to find relevant data assets quickly and understand their context. This includes data source, business or technical definitions, and quality. Data catalogs also facilitate data monetization, regulatory compliance, and self-service analytics.

Data curation starts by defining several elements. Understanding these elements is an important function of the data product lifecycle.

## Data domains

A data domain is the logical grouping of items or areas of interest in your organization. Examples include customer, product, line of business transactions, and so on. These domains support the structure of your data catalog. They can also help guide you through decisions around constructing a thoughtful set of data policies to govern data access and use.

## Data taxonomy

A data taxonomy is the result of classifying business terminology into hierarchical groups to create structure and standardization. The curator defines each term and establishes relationships between it and other associated terms to effectively tag each data asset, making it easier for data consumers to find what they need.

## Metadata and ETL connectors

Connectors ingest data automatically from a wide range of systems and tools, including databases; extract, transform, and load (ETL) tools; and reporting systems. When defining connectors, curators should understand the breadth of their supported toolsets. They should also consider how to ingest business glossaries because they are often kept in spreadsheets or web pages.

If the process of documenting business definitions is not yet underway, introducing catalog functionality might help with the decision of how to ingest relevant business metadata. Some tools support the entry, review, and approval of those definitions.

With these curated elements, those who consume data products can find what they need and understand it quickly, supported by an end-to-end data strategy.

# What are the elements of an end-to-end data strategy?

Data products are a virtuous cycle, curating data enrichment. They can be visualized by analytics and business intelligence tools, monetized in marketplaces, and integrated with other data sources to increase value. Mature data products are maven of insight and a currency used to define and evaluate resourcing, investment, and innovation. How you govern, secure, collect, transform, store, and analyze your data should focus on one goal: supporting data as a product.

Every organization is different, and a one-size-fits-all approach to data productization will lead to compromises that might not be in the best interests of your organization. Data productization is a journey that is neither linear nor incremental and requires contemplating immediate and long-term goals.

To benefit from data productization, you need to understand where you are on the journey. Identify gaps and then create a continuous improvement roadmap for working with data products at both the domain and organization level. The elements of an end-to-end data strategy can help you find the interconnectivity of data products across your organization.

## Defining the structure of the strategy

Structurally, the data and analytics solutions that consume data products should be comprehensive, integrated, and governed. They need to empower broad and deep tools and integrations. At the same time, they must employ domain and line of business processes and policies to secure and protect data. This means that the underlying structures must be highly performant. They should use optimized, managed, and serverless services where possible.

In this structure, the elements of an end-to-end data strategy coalesce across the functional stack. Intentional use of standards-based, open data formats helps prevent proprietary lock-in. This simplifies data operations that surround data lake and lake house structures.

## The characteristics of data and analytics solutions created from data products

### Comprehensive

Equipped with the right tools, with optimal price performance for any user, use case, and data type.

### Integrated

The ability to integrate data that is stored and analyzed in different tools and systems to gain a better understanding of your business and predict what will happen.

### Governed

Governance across the data lifecycle to speed up the pace of innovation.

# Foundational layers of data products

Cloud data lakes, lake houses, warehouses, purpose-built databases, analytics, and serverless storage solutions optimize organizational resources and functionality while reducing undifferentiated heavy lifting. Robust analytical data platforms can consolidate, organize, and democratize data storage. They can also integrate, transform, and manage data pipelines. All these elements are foundational layers of data products. They are extended by machine learning and underpinned by data governance functionality.

Data governance and security is a data productization layer on its own, but it also runs across all the other layers. It is the collection of people, processes, and technology used to ensure data quality and protection throughout its lifecycle. Data curators and stewards play a major role in data governance and security.

Greater organizational alignment driving data product improvement and adoption is critical regardless of the tools and infrastructure you currently have in place. With intentional planning and thoughtful implementation for short and long-term data product goals, every organization can enhance and empower their data-driven outcomes.

Focusing on and reviewing each of these foundational layers can help you identify gaps in your processes and find AWS Partner solutions that work to address them. AWS Partners have a comprehensive set of services and solutions to activate the journey to improved data productization.



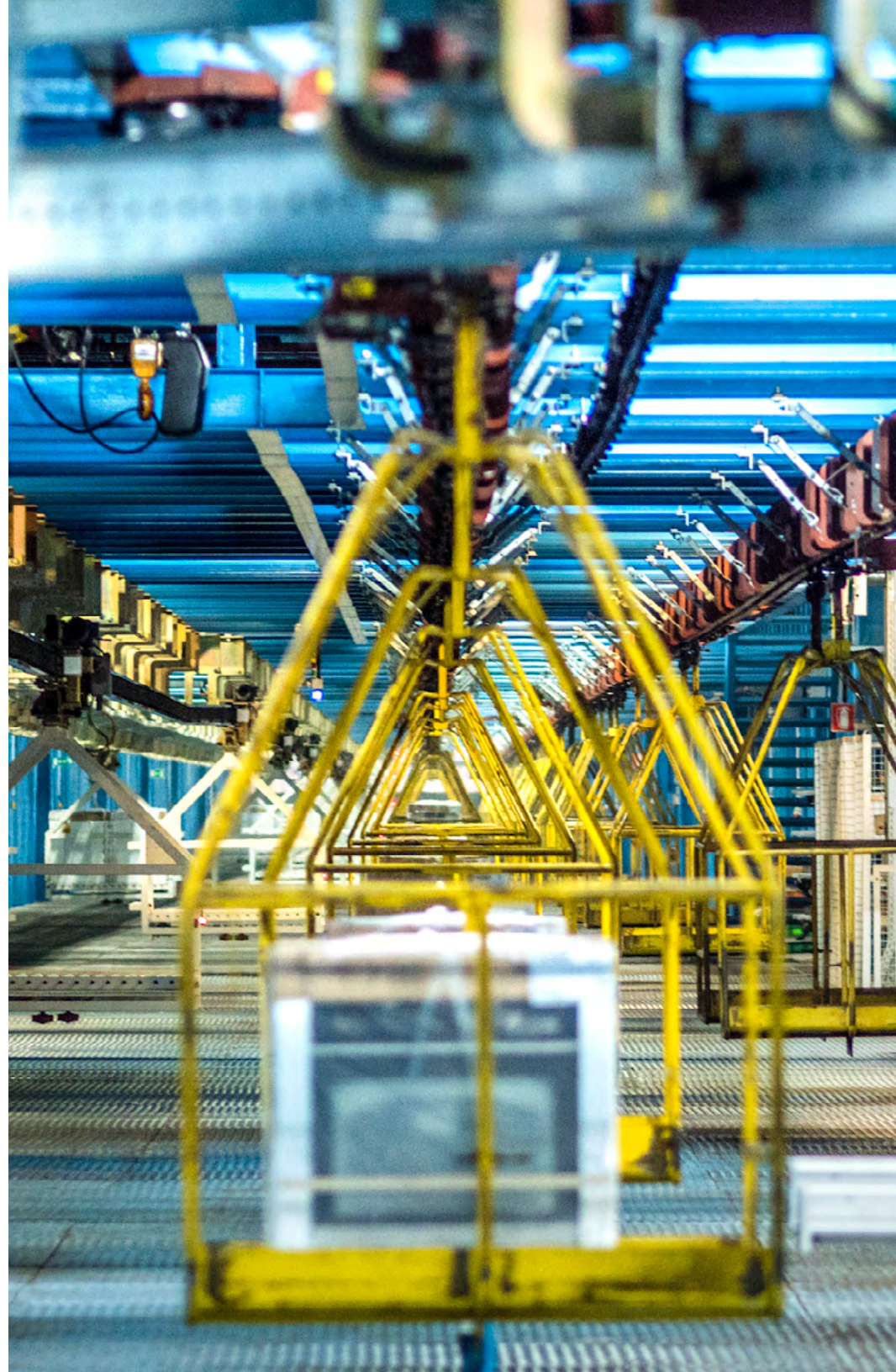
# Unlock the value of your data with AWS and AWS Partners

Every organization has different data sources, analytics needs, and governance requirements. And these requirements are dynamic and change over time. Harnessing data to its full potential requires more than having just one data store or one data lake. Delivering a complete solution for storing, accessing, analyzing, and visualizing data—plus predicting outcomes—is a journey of continuous improvement.

AWS Partner solutions can help you implement an end-to-end data strategy that can handle data scale at any level. Powered by AWS, partner solutions can improve and optimize workloads, data types, and desired business outcomes. You can modernize your data infrastructure from a legacy solution to a scalable, trusted, and secure cloud provider. You can also break down silos, so data can be put to work effectively across databases, data lakes, analytics, and ML services.

AWS Partner solutions enable you to provide personalized customer experiences, build related digital products, optimize operations, and increase employee productivity. Examples include services and platforms that deliver analytics in the format you need. If you want to invent new experiences and data products, you can do that, too.

Partner solutions built on AWS provide what you need to reimagine processes, using purpose-built databases, advanced analytics, visualizations, and AI/ML. AWS Partner solutions provide powerful and widely accessible tools that facilitate data democratization, enabling less technical resources to curate and broker data access and value.



# Conclusion: From data sets to data products

An end-to-end data strategy for data productization delivers the elements and governance that enable a data-driven enterprise. Decisions are made based on data, not gut feelings, or organizational knowledge. So, when you're defining a data product, you need to be intentional about the strategy, the user, the value, the technology, the usability, and how you're going to communicate, provide, and secure access.



## Define your data product goals

Start by examining how your data products will enable users and how these products will align with your business goals. For example, do you want to use business insights or analytics to make strategic recommendations that can increase retail sales? Do you want to monetize your clean and high-quality labeled datasets? Answering these kinds of questions gives you the opportunity to think about the type of value your data offers.



## Consider user needs

Work with your users to learn how they work and what they want. Finance teams may prefer spreadsheet-like applications. Business users may opt for dashboards. Data scientists often start with raw data. Document your findings. Use definitions and language that people understand and use every day.



## Track value with data

Identify the KPIs that show business value. Do you want to reduce costs, improve efficiency, innovate, collaborate, or monetize? Make sure these KPIs are measurable, show tangible outcomes, and support continuous improvement.



## Differentiate on the technology

Technology enables you to create a better data product. AWS Partners collaborate strategically with AWS to deliver cloud services and technology. The goal is to help you take the steps needed to productize your data and use it to generate new revenue streams or new insights.

## Want more information?

The process of data productization starts at the sources of all your data and ends with monetization or data-driven outcomes. It's important that each step in your data productization journey supports improving data as a product.

Focusing on and reviewing each of these steps and how they facilitate data productization can help you identify gaps in your processes. Then you can find AWS Partner solutions that work to address them. The result? Mature data products that define and evaluate resourcing, investment, and innovation.

**Learn more about how your business can unlock the value of your data with AWS >**

