Six best practices for building and deploying effective data products

Table of contents

Data products: How organizations make the most of their data

Organizations today understand just how important and valuable their data is. They are also aware that they need to use data more effectively to achieve better business outcomes. Unfortunately, they also face the challenge of maximizing the usefulness of data assets and improving the efficiency of data workers as the volume, diversity and speed of data continue to grow.

Data products can help.

Data products help organizations bundle data in an easy-to-use format. This enables data consumers of all levels to use data to make better decisions. Using data-driven insights derived from data products allows companies to automate processes to reduce costs, create personalized experiences for customers and identify market trends faster than competitors.

In this ebook, we'll share the best practices for building and deploying data products to achieve business goals, improve decision-making and develop new ideas.

Whether you are new to data products or want to do more with them, this guide provides the practical advice you need to succeed.
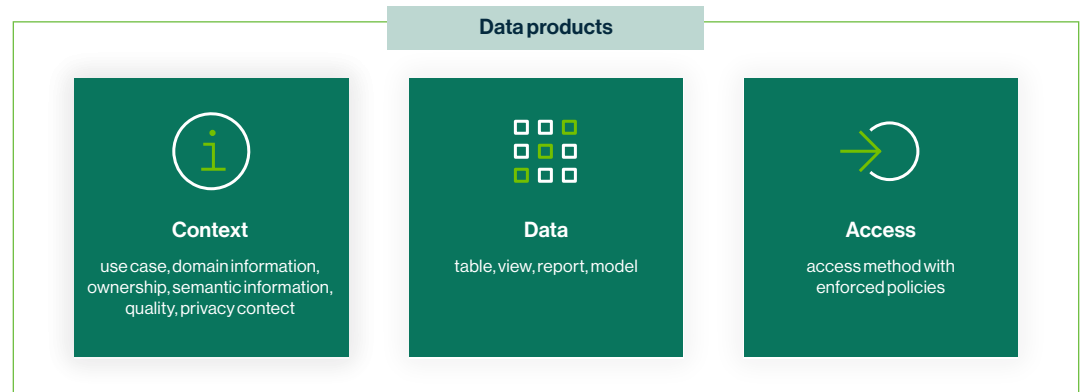
What are data products?

There are several definitions of the term "data products." Essentially, data products are reusable packages that contain the underlying data, tools, documentation and context required to access, understand and utilize the data to achieve specific business goals. Data products include data pipelines, schemas, machine learning models, user interfaces, governance policies and access mechanisms.

The composition of data products can vary, and they can be refined and combined to create new data products. They can range from general purpose to highly customized, depending on the specific functionality required to meet business objectives.

Key components

- **Context:** Information needed to understand and use the data, including its business use case

- **Data:** The main information in the product

- **Access:** Ways to get and work with the data

Data products

Context	Data	Access
use case, domain information, ownership, semantic information, quality, privacy contect	table, view, report, model	access method with enforced policies

Examples

- **Sales dashboard:** Provides near real-time insights into sales performance with visualizations and filters

- **Customer segmentation dataset:** A curated dataset identifying customer groups for targeted marketing

- **Fraud detection model:** An AI model that flags suspicious transactions based on historical data

Data-as-a-product vs. data products

Using product management principles to enhance the use and value of data is known as data-as-a-product. It involves ensuring that it's built, managed and delivered, focusing on the end-user's needs.

Data products are the output of treating data as a product. They are packaged entities that include data and other contents such as relevant metadata, pipelines and documentation, resulting in valuable curated data assets.

Although the data-as-product approach and data products are commonly associated with the data mesh methodology, not all organizations need data mesh to benefit from these concepts.

Why data products?

Data products are reusable, complete packages containing the underlying data, tools, documentation and context that users need to understand, access and utilize data. When organizations want to meet specific objectives using their most critical data — whether it's driving strategic decisions, improving operations or enhancing customer experience — data products offer an ideal way to achieve specific business goals:

Align data management efforts with business goals and AI-readiness

Data products ensure that data management supports business objectives and AI initiatives. They provide high-quality, compliant and relevant data, improving decision-making, AI model performance and business outcomes.

Empower self-service access, data democratization and innovation

The IDC December 2022 Data Valuation Survey found that almost 44% of the 1,024 respondents cited that they don't have access to the data they need to do their jobs. By making high-quality data available to various teams, data products break down departmental silos, promote a data-driven culture and support faster decision-making. Accessible data encourages creativity and innovation, allowing teams to develop new analytics, applications or services and respond quickly to market changes.

Improve productivity while reducing costs and duplication

Untimely data delivery was a major data pain point for 33% of executives surveyed in a recent MIT Technology Report. Well-defined data products provide stakeholders with accurate and timely information, enabling them to make informed decisions faster to improve business outcomes. Additionally, by providing reusable data assets, data products save time and resources by eliminating the need to recreate data assets for different uses, maximizing the return on data investments.

Drive accountability and compliance

Domain ownership and structured frameworks for data products help ensure responsibility for data quality and security, which encourages responsible data management. Managed data products also help ensure that governance policies and regulations are followed, reducing the risk of non-compliance.

6 best practices for integrating data products in your organization

1. Have a clearly defined purpose and value

Why are we doing this?

Begin by clearly articulating the purpose of the data product. What specific problem or business question is it designed to address? Can you define measurable goals to gauge success? It's important to be able to define your business reasons and frame each goal in terms of potential benefit to your organization:

- **Competitive advantage:** How will developing this data product (or data products) create new opportunities?

- **Improved decision-making:** How can providing accurate and timely information to stakeholders enable them to make informed decisions and improve business outcomes?

- **Cost reduction and avoiding duplication:** How do reusable data products save time and resources by eliminating the need to recreate datasets for different uses and maximizing the return on data investments?

2. Focus on usability

Who is our target audience? How will they use the data product?

Identify who will use the data product and how they will interact with it. Design with the end-user in mind to ensure the product is intuitive and meets their needs.

For example:

- **Customized executive dashboards for executives and senior management** can provide near real-time visualizations of critical business metrics. These dashboards are often built using cloud-based business intelligence (BI) tools to ensure accessibility from any device, including mobile. The dashboards (which use advanced security protocols to safeguard sensitive company data) are typically role-based to ensure executives can see insights tailored to their scope supporting strategic decision-making.

- **Marketing teams can use web-based campaign performance dashboards** to track the success of their digital and offline campaigns. These dashboards, which can be shared across teams using adjustable access privileges to protect sensitive marketing budget data, monitor critical metrics such as click-through rates (CTR), conversion rates, cost per acquisition (CPA) and return on investment (ROI). With API-enabled automated data feeds and consolidated reporting from marketing platforms (e.g., Google Analytics, social media ad managers), the dashboards offer near real-time updates, enabling quick adjustments to optimize ongoing campaigns.

- **Financial analysts can leverage risk management models** that apply advanced predictive analytics and scenario modeling to evaluate financial risks. These tools, accessed through secure platforms that integrate with enterprise financial systems, provide insights into potential market fluctuations and credit and operational risks, which help analysts build risk mitigation strategies. The models incorporate external data sources (which may contain attributes that change periodically), such as economic indicators and industry trends, to enhance the accuracy of forecasts. Analysts can access reports via dashboards or download data for further analysis in a spreadsheet or analytics software, ensuring flexibility in how they use the data.

3. Optimize discoverability and accessibility

How will end-users discover this data product? How can you drive adoption?

To ensure users can easily find and utilize your data product, focus on discoverability and accessibility. Hosting the data product in a well-organized repository for curated data assets — like a data marketplace — is a best practice that simplifies access for users of all levels. Make the data product easily searchable by incorporating relevant metadata and tags to enhance visibility during browsing or searches within the platform.

An intuitive user interface (UI) is crucial for driving adoption. The design should be user-friendly, featuring clear descriptions, usage examples and well-labeled methods for accessing the data product to improve the overall experience.

Incorporating user feedback mechanisms like ratings and reviews helps build trust and encourages engagement. Additionally, ensuring that your data product complies with governance and privacy standards is vital—users are more likely to adopt solutions prioritizing security and compliance.

How do you access data products?

A data marketplace is a centralized hub where users can discover, assess and access trusted data products. It facilitates better organization and searchability of data assets, making it easier for data consumers to find the data they need.

4. Ensure reliability and trustworthiness

What's the best way to assure end-users about the quality of the data product and its underlying data?

Demonstrating the reliability and trustworthiness of data products is essential for building end-user confidence. Here are some practical ways to achieve that:

- **Provide comprehensive documentation** that details the source data, and trace and communicate the [data's lineage](#), explaining where it came from and any transformations applied. This enhances transparency and allows users to assess the data's origin and journey, building trust and helping users understand the data's context.

- **Display data quality metrics** such as accuracy, completeness, consistency, timeliness and validity. Presenting these metrics can help users evaluate the data's reliability.

- **Certify data products** to endorse their reliability and readiness for use. Organizations can tailor the certification process to meet their requirements, enabling data consumers to easily identify certified assets through visible labels. This endorsement indicates that the asset meets established standards and instills confidence in users, ensuring they can trust the data for informed decision-making.

- **Enable crowdsourced feedback** where users can leave reviews and ratings for the data product. Positive testimonials and user experiences can significantly enhance perceived quality and trustworthiness.

What is a data contract?

A data contract is a vital element of a data product, representing the promise made by the data owner to data consumers. It outlines expectations, including service-level objectives (SLOs) related to system uptime and latency for a data product. Additionally, it details relevant information about pipelines or data delivery mechanisms and can also provide information about the data itself, such as schema and expected quality metrics.

This distinction highlights the difference between a data contract – which ensures the data is suitable for the consumer's needs – and a data sharing agreement, which defines allowed and prohibited uses of the data. Together, they create a balance between maximizing the value derived from data and protecting data from potential misuse.

5. Ensure that the data product meets integration and interoperability standards

### Does the product meet organizational standards for interoperability?

The data product must meet integration and interoperability standards to optimize data flow and operational efficiency across different systems and departments. To achieve this, well-known data standards such as JSON should be used to establish consistency standards for internal data. Forming a governance committee, appointing data stewards and employing data integration platforms and APIs are critical to ensuring seamless integration and interoperability.

6. Consider scalability and maintainability

What mechanisms exist to collect feedback? How will maintenance affect end-users?

A robust data governance framework is essential for ensuring that data products maintain compliance and consistency as they scale. This framework helps organizations define clear policies and standards, which not only enhance data quality but also build trust among users. Also, best practices include defining lifecycle stages like 1) defining a business use case, 2) identifying and understanding data that will be used, 3) documenting the data product, and 4) verifying and monitoring the data product.

Data product framework
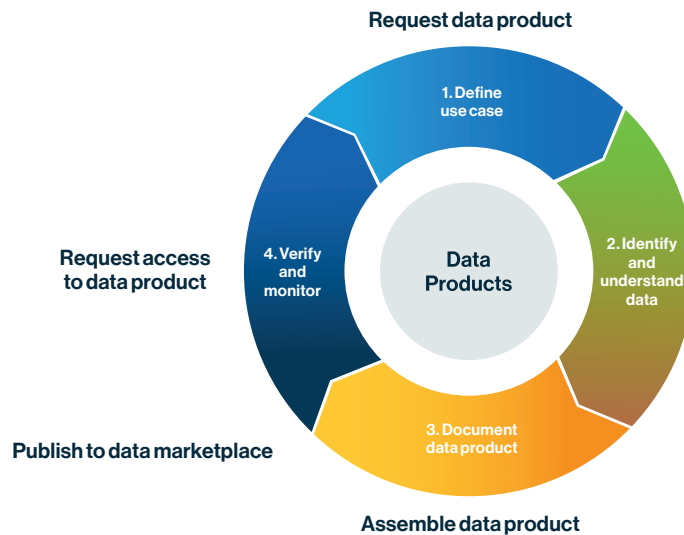
Additionally, seamless integration capabilities allow for easy connectivity with a wide range of data sources and tools, enabling organizations to build data products more efficiently. Coupled with automated workflows, which streamline data management processes, organizations can minimize the manual effort needed to maintain and scale data products. This combination of governance, integration and automation empowers organizations to adapt to changing data needs efficiently while ensuring high standards of quality and compliance.

Engaging users early and using an iterative development approach is vital for the successful maintenance of data products. This can include collecting feedback continuously through ratings, reviews and product analytics, for example.

Critical capabilities for managing data products

How can you ensure the data product does what it needs to do?

Managing data products requires flexibility, automation and user-friendly tools to help organizations get the most value from their data. Here are some essential capabilities for building and deploying data products:

Flexible data product framework

A flexible framework is crucial for organizations to deploy data products based on their needs while following industry standards and best practices. Pre-built, out-of-the-box (OOTB) solutions can provide useful templates and guidance but should also be customizable. Organizations should be able to define their data product components — such as business use cases or data contracts — and lifecycle stages.

This flexibility makes it easier to curate, enrich and publish data products, and streamlines the development process. Automation further helps by simplifying discovery, metadata enrichment and publishing to a data marketplace, ensuring products are readily accessible and up-to-date.

Seamless access and collaboration

Easy access and collaboration are essential to make the most of data products. A robust data marketplace acts as a central hub where data consumers can easily search for and discover relevant data. Features like semantic search, filters and data similarity-driven suggestions help users find what they need and discover other relevant assets. Contextual data asset views provide deeper insights into data products, helping users assess their usefulness through previews and detailed asset pages. Allowing consumers to request access, review comments and ask questions, enhances collaboration and ensures smooth data sharing between teams.

Automated data access management

Managing who can access which data product is a crucial part of governance. Data sharing agreements and access management protocols should be in place to ensure the right people have access to the right data while complying with data privacy and protection standards. Flexible options for request fulfillment, whether automated through workflows or integrated with tools like Jira, ensure timely access and delivery.

Additionally, the ability to query data within data products allows users to explore the underlying data and derive meaningful insights.

Usage tracking and reporting

Tracking data product usage provides valuable insights into their adoption and value within an organization. Analytics and monitoring tools help users understand which data products are most popular or valuable and highlight trends in user behavior. This data can guide future development efforts, product prioritization and continuous improvements, fostering a data-driven culture.

These critical capabilities ensure that organizations can build and deploy data products that are easily discoverable, understandable, trustworthy, accessible and valuable for the appropriate end-users.

Unleashing data's full potential with data products

By now, you should understand why building and deploying effective data products is essential for maximizing data's value within an organization. By following these best practices, you can ensure that your data products deliver value by driving innovation and aligning with business goals.

You should also understand more about how a data marketplace is the perfect way to provide access to curated data products. And how protecting data products from the start ensures greater usability and compliance.

Embrace the principles outlined in this guide to create data products that meet current needs and adapt to future challenges and opportunities.

Get started with data products
Learn more about how data products can help you deliver value, drive innovation and realize business objectives here.

1. IDC PlanScape: Data as a Product, Stewart Bond, Chandana Gopal, Lynne Schneider, May 2024
2. Modernizing Data with strategic purpose, MIT Technology Review Insights, April 2024