DataArt

Unlocking Business Value with Data as a Product: A People, Process, and Technology Perspective

DataArt

In today's data-driven world,

organizations recognize data as a valuable asset. However, merely possessing data isn't enough. The real potential lies in effectively managing, accessing, and leveraging it to drive meaningful business outcomes. Traditional data management approaches—centered around IT-driven projects and static dashboards—often fall short due to long time-to-value, limited reusability, and a lack of scalability. Also, as they often have a heavy bias on data producer perspective, i.e., here is the data I have, and lack deep understanding of the consumers – what business problems and objectives they have and how data and actionable insights can really help to address those.

This paper advocates for a paradigm shift: adopting a "Data as a Product" (DaaP) —a structured approach to building reusable, high-quality data products that empower business users, accelerate time-to-insight, and drive significant business value. It examines this concept through the lens of people, processes, and technology, exploring current trends, value outcomes, potential challenges, and strategies to overcome them.

What is Data as a Product?

Data as a Product treats data like any other product—packaged and delivered. It emphasizes usability, discoverability, maintainability, and, crucially, evolution. It allows for products to be easily accessed and consumed in a standardized and self-described way, eliminating the need for several lengthy interactions with the producer. Unlike ad-hoc data projects, DaaP creates reusable data assets consumable by multiple stakeholders across the organization. This involves defining clear ownership, establishing data quality standards, and providing easy-to-use interfaces for accessing and understanding the data (see Figure 1).
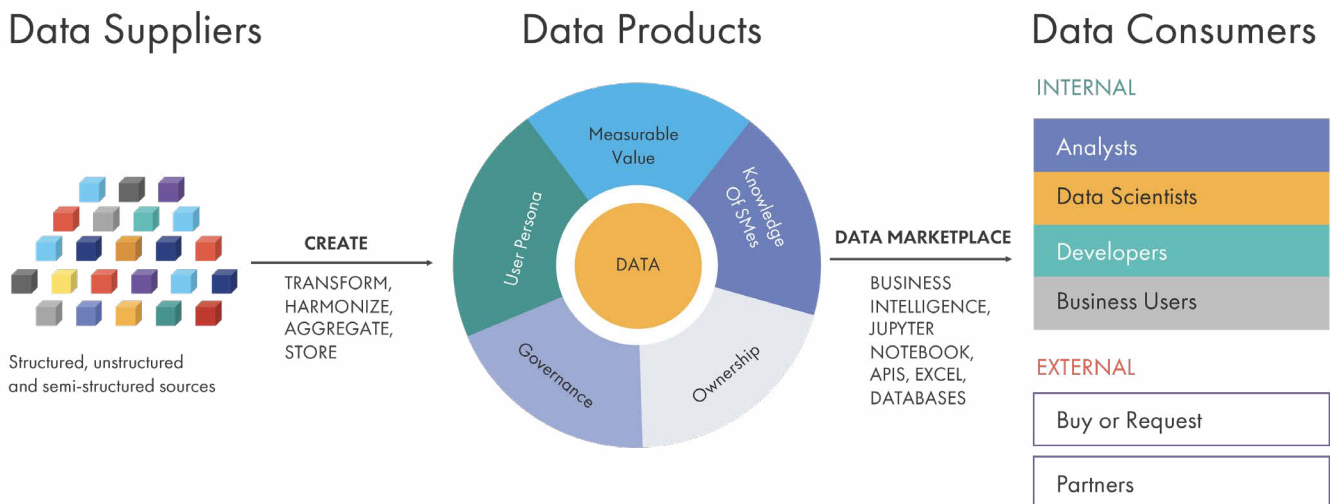

Figure 1: Creating Value for Data Consumers with Data Products

The Role of a Data Product Owner

Like any product, a data product has a lifecycle. It starts with an initial release, undergoes iterative improvements based on user feedback and evolving business needs, and may eventually be deprecated. This iterative nature necessitates strong data product ownership. A dedicated data product owner **is responsible for**:

Defining the product vision and roadmap: Aligning the data product with business objectives and prioritizing features based on user needs and market trends.

Managing the product backlog: Prioritizing data product evolution based on consumer use cases and their business value, integrating new data sources based on availability and quality, addressing data quality issues, and managing technical debt.

Gathering user feedback: Understanding how the data product is being used and identifying areas for improvement.

Managing the product lifecycle: Overseeing releases, updates, and eventual deprecation of the data product or some of its features.

Leading the data product engineering and operations team: Working with the engineering team to implement and evolve the data product and with the data operations team to ensure the operational aspects, such as reliability, observability, and responsiveness to user requests, issues, and concerns are timely and proactively addressed.

Balancing data supplier availability and consumer needs: Acting as the bridge between data suppliers (providers of raw data or source data products) and data consumers (business users utilizing the data product), they balance the availability and quality of source data with the specific needs and use cases of the consumers, ensuring the product delivers maximum value. This includes negotiating SLAs with data suppliers and advocating for improvements in data quality or availability when necessary. Increasingly often, data product and software engineering may be handled by a single domain-aligned unit with tight collaboration. In such cases, the data product owner can play an active role in influencing software product requirements to ensure that the needs of the data product consumers are built into the design of the source software systems.

This product-centric approach ensures that data is not just collected and stored but also actively managed and refined to deliver maximum business value.

The Business Impact of Data as a Product:

The DaaP approach offers several **key business benefits:**

Increased Value: The key objective of the DaaP is to maximize the data product value based on deep and continuous understanding of the user personals, their business objectives and challenges. Instead of dumping piles of data on them and hoping they will make some use of it, data products go much further in making sure they address real needs.

Accelerated Time-to-Value: Pre-built, readily available data products allow businesses to quickly access information for informed decisions, significantly reducing data gathering and preparation time. Additionally, a self-service data platform enables data product engineering teams to build, integrate, and deploy data products without the need to develop separate infrastructure for each one. This accelerates the data product

team's workflow, allowing them to focus entirely on delivering the data product itself.

Increased Data Democratization: DaaP empowers business users to access and utilize data without heavy reliance on IT or data teams, fostering a data-driven culture and enabling faster, more agile decision-making.

Improved Data Quality, Consistency, and Governance: Top-down broad organizational data governance programs often fail to make a significant positive impact or real quality, accessibility, and utility of data and mostly focus on policies and responsibilities. Data as a Product brings very specific value-driven motivation to data governance and quality in the scope of every specific data product, as the impact of lack of quality, consistency, or security can be clearly seen, while good practices lead to more reliable insights and better business outcomes.

Enhanced Innovation and Agility: Easy access to high-quality data fosters innovation and allows businesses to quickly adapt to changing market conditions.

Reduced Costs and Increased Efficiency: Utilization of self-service data platform across the entire landscape of organizational data products simplifies infrastructure, reduces overall effort and cost, and standardizes capabilities, such as data accessibility, observability, security, monitoring, and so on. Reusing existing data products reduces redundant data projects, leading to significant cost savings and increased efficiency.

The People, Process, and Technology Triad: A Data Factory Approach

Implementing a successful DaaP strategy requires a holistic approach considering the interplay of people, processes, and technology.

1. People: Building a Data-Driven Culture

Challenge: Siloed organizational structures, lack of data literacy, and unclear roles and responsibilities can hinder DaaP adoption.

Solutions:

Establish Data Product and Data Ownership: Clearly define roles and responsibilities for the data product owners, as well as data owners, stewards, and consumers.

DataArt

Foster Data Literacy: Invest in training and education programs to improve data literacy across the organization.

Create a Data-Driven Culture: Encourage data sharing, collaboration, and experimentation.

Establish Cross-Functional Data Product Teams: These teams should include data engineers, data scientists, business analysts, and subject matter experts, working together to define, develop, and maintain the data products.

Set Up Data Platform and Data Ops Teams: The data platform team builds and evolves a data platform to enable data product teams' self-service capabilities. The data ops team looks after operational capabilities, such as reliability, observability, user issues, and automation to better, faster, and cheaper serve the data consumers.

2. Process: Standardizing Data Product Management

Challenge: Lack of standardized processes for data product development, maintenance, and governance can lead to inconsistencies and inefficiencies.

Solutions:

Establish a Data Product Lifecycle: Define, develop, test, deploy, and maintain data products.

Implement Data Governance Frameworks: Establish clear organizational data quality standards, data security policies, and data access controls and set of best practices to embed those into each data product lifecycle.

Apply Product Thinking and Product Design: Use the product design principles to focus on data consumer engagement to deeply understand the needs, value, and their business context.

Embrace Agile Methodologies: Use agile principles to iterate quickly and adapt to changing business needs.

Establish Clear Metrics and KPIs: Define how the success, impact, and value of the data products will be continuously measured and monitored.

3. Technology: Enabling Scalable and Accessible Data Products

Challenge: Legacy systems, data silos, and a lack of appropriate tools can hinder the creation and delivery of data products.

Solutions:

Self-Service Data Platform: The data platform that brings all needed technology and horizontal capabilities, such as data storage, processing, security, or observability, and allows data product teams to build and deliver data products by focusing on business aspects, data pipelines, and models. The platform will also allow end-users to discover, access, and consume data products in a self-service mode.

Modern Data Stack: Utilize a modern data stack that includes cloud-based data warehouses, lakehouses, meshes, fabrics, technologies for data pipelines, data CI/CD, data quality, observability, security, metadata, and data catalog and governance tools.

Data Catalog, Semantic Layer, and Metadata Management: Implement a robust data shopfront to enable data discovery, understanding, accessibility, and governance. This layer may include a data catalog and semantic layer and should be augmented with AI to automate metadata enrichment, data profiling, and data lineage tracking.

Data Contracts and API-First Approach: Standardized data contracts and data product access APIs are a key technology pillar of the DaaP paradigm. Data contracts capture everything that is needed to use a data product, such as data schema and semantics, quality, ownership, SLA, licensing, and pricing. Consider adopting **Open Data Product Specification** or **Open Data Contract Standard** to enable data product consumption and interoperability beyond your organization.

Data Observability: Implement data observability tools to monitor data quality, performance, and usage; prevent or quickly investigate and address data issues across your entire data landscape.

Organizational Model to Succeed in Becoming Data-Driven

To succeed in a data product-driven organizational model, alignment between foundational data capabilities and business-focused data products is essential. This requires a two-tiered structure: Horizontal Data Platform Foundation Teams and Vertical Domain

Data Product Teams (see Figure 2). The foundation teams build and maintain scalable data infrastructure—handling ingestion, transformation, storage, governance, and security—ensuring a robust ecosystem for data products. Vertical teams, aligned with business domains, transform raw data into curated, trusted, and explainable products. This structured approach enables organizations to maximize data's potential while ensuring agility, trust, and usability.
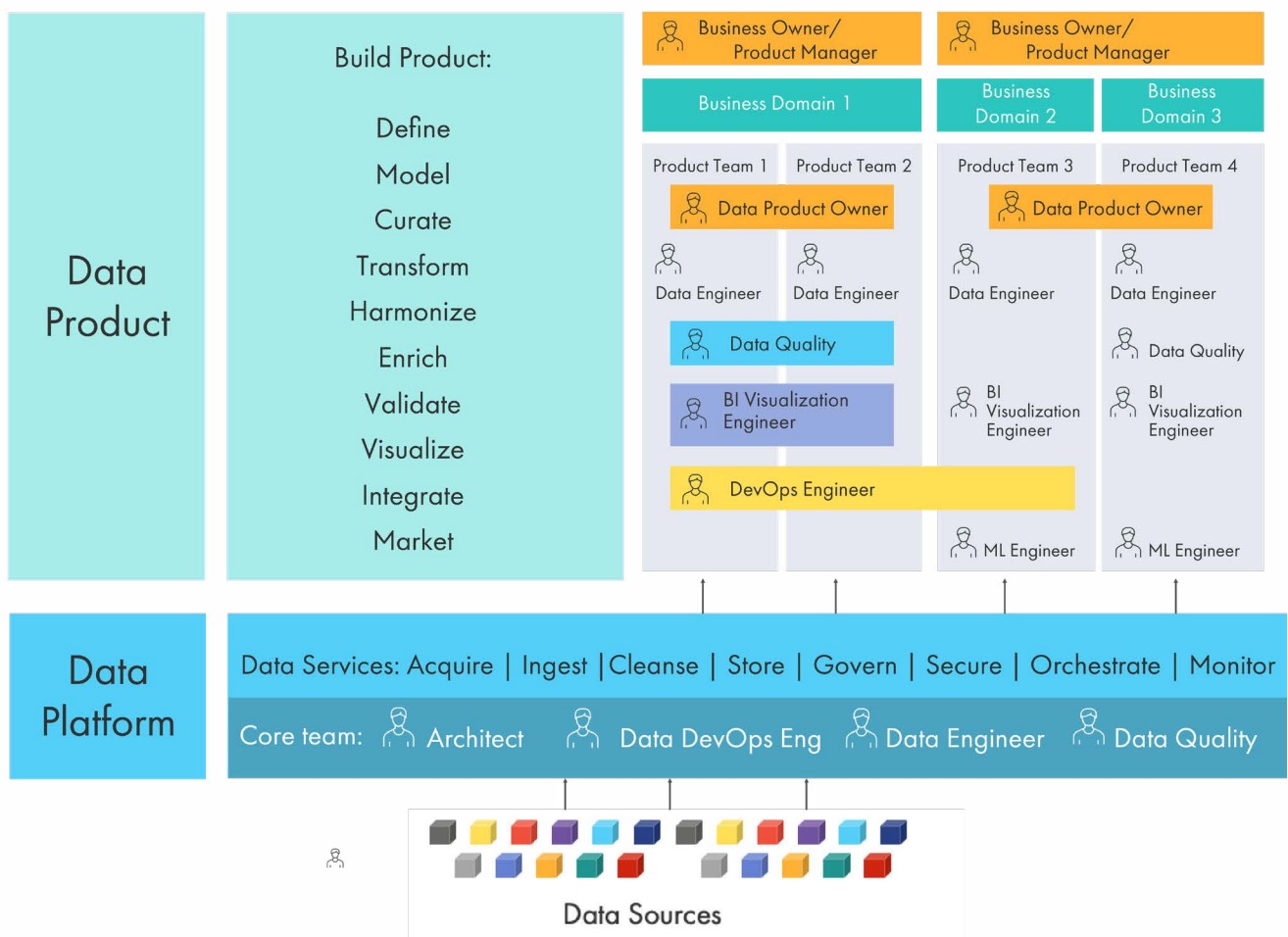

Figure 2: Two-Tiered Organizational Structure in Becoming Data-Driven

Horizontal Data Platform Foundation Team(s):

These teams build and maintain the data capabilities of the organization, focusing on creating a robust and scalable data platform, a foundation for building data products, including:

Data Ingestion, Data Pipelines, Data Modeling, and Transformation: Integrating

technologies and providing blueprints for building efficient and reliable data pipelines to ingest data from various sources and to consume source data products; then, model, enrich, and transform the data into the data product end-state. Data quality, metadata, lineage, security, orchestration, monitoring, versioning, CI/CD, and other horizontal aspects also need to be addressed in these data pipeline technologies and blueprints.

Data Storage and Processing: Implementing a scalable and cost-effective data warehouse, lake, lakehouse, mesh, or fabric architecture to store data, run pipelines, and serve data products to consumers.

Data Catalog, Metadata Management, Semantic Layer: Consumer-facing data shopfront allowing to discover, understand, and access data products and all the metadata foundations that need to enable this, as well as support data engineers.

Data Observability: Implementing tools and processes to monitor data quality, lineage, and performance, ensuring data reliability and trustworthiness.

Data Security and Governance: Implementing security measures and governance frameworks to protect sensitive data and ensure regulatory compliance.

These horizontal teams provide the raw materials and infrastructure upon which the vertical product teams build.

Vertical Domain Data Product Teams:

These teams are aligned by specific business domains or data product verticals (e.g., Customer 360, Product Analytics, Supply Chain Optimization). They act as "product factories," taking the raw materials from horizontal teams and transforming them into curated, trusted, and explainable data products.

Example: Customer 360 as a Data Product:

A common DaaP example is a "Customer 360" product. Instead of building separate customer dashboards for different departments, a single, comprehensive customer data product is created. This product combines data from various sources (CRM, marketing automation, e-commerce, etc.) to provide a unified customer view. This product can then be consumed by marketing, sales, customer service, and other departments, ensuring consistent and accurate customer insights.

The Data Product Teams responsibilities include:

Data Curation and Transformation: Cleaning, transforming, and enriching raw data

to meet the specific requirements of the data product.

Data Product Development: Building and maintaining the data product, including defining data models, creating APIs, and developing documentation.

Data Quality Assurance: Ensuring data product quality and accuracy through rigorous testing and validation.

Explainability and Transparency: Ensuring the data product is understandable and explainable to business users, fostering trust and adoption.

Delivering Actionable Insights: The ultimate goal is to create data products that facilitate decision-making, provide valuable insights, and trigger next-best-action recommendations. This division of responsibilities between these teams brings required focus to both data platform and data products and ensures efficient building and maintenance of foundational data capabilities while allowing specialized expertise in developing and delivering valuable data products. This structure directly supports the DaaP philosophy by establishing clear ownership and accountability for each data product.

As organizations evolve toward a data-driven future, Artificial Intelligence (AI) is becoming a powerful enabler of Data Products—enhancing their creation, integration, and impact. From automating data preparation to embedding intelligence within products and driving smarter decision-making, AI is redefining how businesses extract value from data.

Unlocking the Potential of AI in Data Products: From Creation to Intelligent Action

As organizations evolve toward a data-driven future, Artificial Intelligence (AI) is becoming a powerful enabler of Data Products—enhancing their creation, integration, and impact. From automating data preparation to embedding intelligence within products and driving smarter decision-making, AI is redefining how businesses extract value from data.

AI in Creating Data Products

Artificial Intelligence is transforming how data products are built by automating critical

DataArt

processes such as data ingestion, cleansing, and transformation. Machine learning models can detect patterns, fill gaps in datasets, and improve data quality, making it easier for organizations to create trusted, high-value data products. Additionally, Generative AI can assist in automated documentation, metadata generation, and schema recommendations, streamlining data product development.

AI as Part of Data Products

AI doesn't just aid in the creation of data products—it can become an integral part of them. AI-powered recommendation engines, predictive analytics, and natural language interfaces allow business users to interact with data products more intuitively. AI-driven personalization also enhances self-service analytics, enabling users to discover the most relevant insights based on their needs and context.

AI and Decision Intelligence with Data Products

AI-driven Decision Intelligence takes data products a step further by turning insights into action. By leveraging real-time data, predictive modeling, and automation, AI can help organizations make faster, smarter, and more proactive decisions. Whether optimizing supply chains, improving customer experience, or detecting anomalies in financial transactions, AI-powered decision intelligence ensures that data products don't just inform but drive meaningful business outcomes.

Challenges and Mitigations

While AI offers immense potential in data products, it also presents challenges:

Data Quality & Bias: AI models are only as good as the data they learn from. Implementing robust data governance and continuous monitoring ensures reliability and fairness.

Security & Privacy Risks: AI-powered data products must comply with regulatory standards (e.g., GDPR, CCPA) and include encryption, anonymization, and access controls.

Organizational Adoption: AI adoption requires a shift in mindset and skills. Training, clear communication, and change management are critical to ensuring that AI-driven data products are trusted and effectively utilized.

 As AI continues to evolve, its role in data products will expand, unlocking new opportu-

nities for efficiency, automation, and intelligent decision-making. Stay tuned for our next thought leadership article, where we explore AI's deeper impact on data-driven organizations.

Conclusion:

The Data as a Product approach represents a fundamental shift in how organizations manage and leverage data. By focusing on people, processes, and technology within a Data Factory organizational structure and embracing the potential of generative AI, businesses can unlock the true value of their data assets, drive innovation, and gain a significant competitive advantage. This approach moves beyond ad-hoc projects and static reports toward a more scalable, reusable, and impactful way of working with data. Organizations that make this shift will be well-positioned for the future of data-driven innovation.