QuantumBlack
AI by McKinsey

How to unlock the full value of data? Manage it like a product

June 14, 2022 | Article

By Veeral Desai, Tim Fountaine, and Kayvaun Rowshankish

Creating reusable data products and patterns for piecing together data technologies enables companies to derive value from data today and tomorrow.

DOWNLOADS

↓ Article (7 pages)

Our recently published article in *Harvard Business Review*, "A better way to put your data to work," details how to establish a sustainable path to value. The key is to manage data just as you would a consumer product. Here, we present a visual summary of this approach.

Today's predominant—and largely unsuccessful—approaches to data

Organizations typically employ either a grassroots or big-bang data strategy—neither of which enables them to make the most of their data investments.

Grassroots approach

In a grassroots approach, individual teams must piece together the data and technologies they need. This approach results in significant duplication of efforts and a tangle of bespoke technology architectures that are costly to build, manage, and maintain.
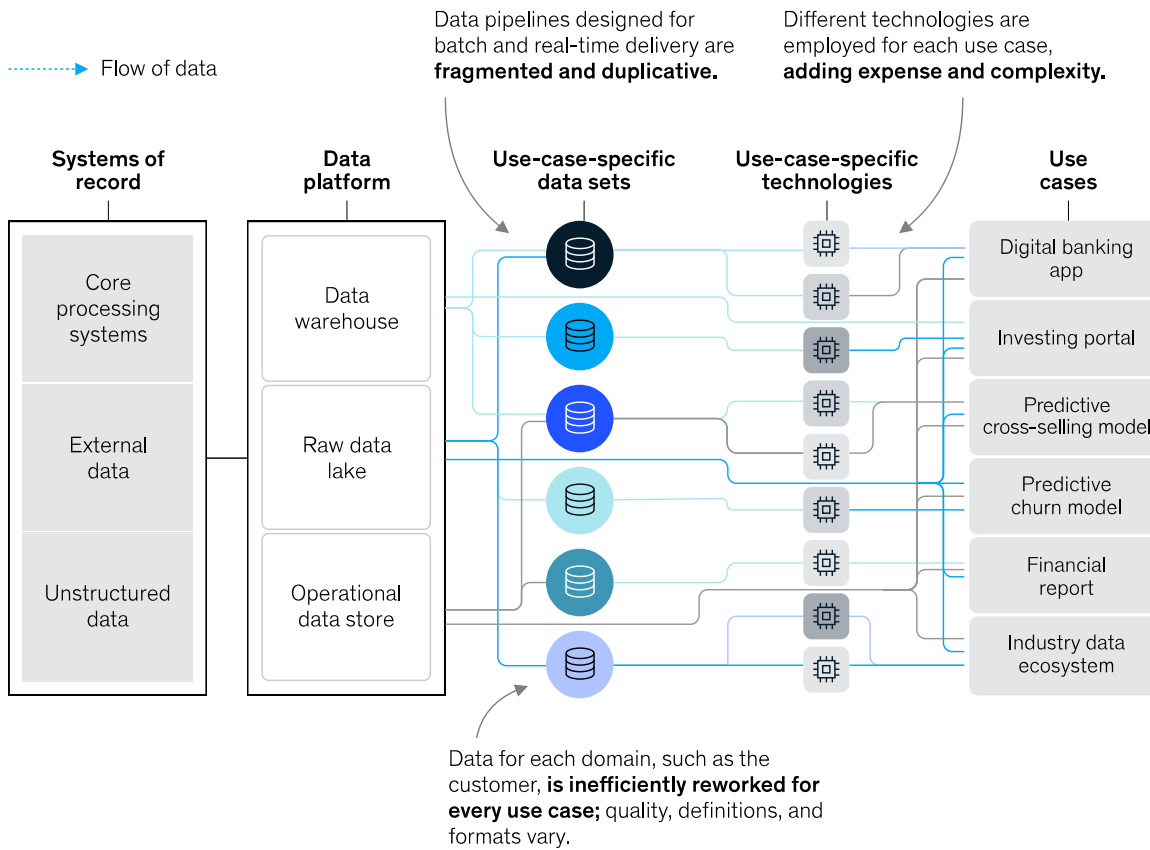
Big-bang strategy

At organizations employing the big-bang strategy, a centralized team extracts, cleanses, and aggregates data en masse. This approach can eliminate some of the rework that occurs, but it's often not aligned with business use cases and therefore fails to support end users' specific needs. End users often struggle to confirm that the data provide the necessary level of governance and quality, which limits the time savings. Later work on new use cases that are aligned with business value often triggers a grassroots approach and its associated problems.

These strategies fail to lay the foundation for current and future use cases that will create value.

Neither the grassroots nor big-bang approach will enable organizations to derive sustainable value from data.

Example of a complex and inefficient data strategy

Data pipelines designed for batch and real-time delivery are **fragmented and duplicative.**

Different technologies are employed for each use case, **adding expense and complexity.**

┄┄┄► Flow of data

| Systems of record | Data platform | Use-case-specific data sets | Use-case-specific technologies | Use cases |

Core processing systems · External data · Unstructured data

Data warehouse · Raw data lake · Operational data store

Digital banking app · Investing portal · Predictive cross-selling model · Predictive churn model · Financial report · Industry data ecosystem

Data for each domain, such as the customer, **is inefficiently reworked for every use case;** quality, definitions, and formats vary.

Source: "A better way to put your data to work," *Harvard Business Review*, June 2022

McKinsey & Company

A better approach: Managing data like a product

We find that when companies instead manage data like a consumer product—be it digital or physical—they can realize near-term value from their data investments *and* pave the way for quickly getting more value tomorrow.

Data products are similar to consumer products in many ways.

Examples of similarities

	Digital product *Example: Computer app*	Physical product *Example: Car*	Data product
Product features			
Customization of base product for different users	App enables users to personalize the layout, color schemes, and content displayed and to select plans and pricing structures that meet their needs	Car buyers may purchase a variety of special options (eg, leather upholstery, tinted windows, antitheft systems)	Data products can be wired to support different systems that consume data, such as advanced analytics or reporting systems
Delivery of regular product enhancements	Automatic downloads of new functionality	New models Engine modifications that boost fuel economy	New data Support for additional consumption archetypes
Production efficiency			
Reuse of existing processes, machinery, and components	Software developers reuse blocks of code	Automakers use a common chassis on vastly different cars	Organizations reuse blueprints and modular technologies for consumption archetypes across products

McKinsey
& Company

Data products provide all the data on one entity

A data product delivers a high-quality, ready-to-use set of data that people across an organization can easily access and apply to different business challenges. For example, a data product could provide a 360-degree view of an important entity, such as customers, employees, product lines, or branches. Or it could deliver a given data capability, such as a digital twin that replicates the operation of real-world assets.

Data products are wired to enable standard types of consumption

Data products incorporate the wiring necessary for different business systems, such as digital apps or reporting systems, to "consume" the data. Each type of business system has its own set of requirements for how data is stored, processed, and managed; we call these "consumption archetypes."

While an organization might have hundreds of use cases on its road map, they typically fit one of five primary consumption archetypes. Data products built to support one or more of these consumption archetypes can easily be applied to multiple business applications with similar archetypes.

Five consumption archetypes describe the primary ways that users consume data.

| Archetype	Requirements	Example uses
Digital applications	Specific data cleaned and stored in a particular format and frequency (eg, delivering access in real time to event streams of GPS or sensor data)	A marketing trends app or vehicle tracking app
Advanced analytics systems	Data cleaned and delivered at a certain frequency and engineered to allow processing by machine learning and AI systems	Simulation and optimization engines
Reporting systems	Highly governed data with clear definitions—managed closely for quality, security, and changes—aggregated at a basic level and delivered in an audited form	Operational or regulatory compliance dashboards
Discovery sandboxes	A combination of raw and aggregated data	Ad hoc analysis for exploring new use cases
External data-sharing systems	Adherence to stringent policies and agreements about where the data sit and processes for managing and securing data	Banking systems that share fraud insights

McKinsey
& Company

Data products enable more speed and efficiency

Teams using data products don't have to waste time searching for data, processing it into the right format, and building bespoke data sets and data pipelines—an effort that ultimately creates an architectural mess and governance challenges.

A data product approach results in standardization that saves time and money.

Example of an efficient data product approach

- - - - - ▶ Flow of data
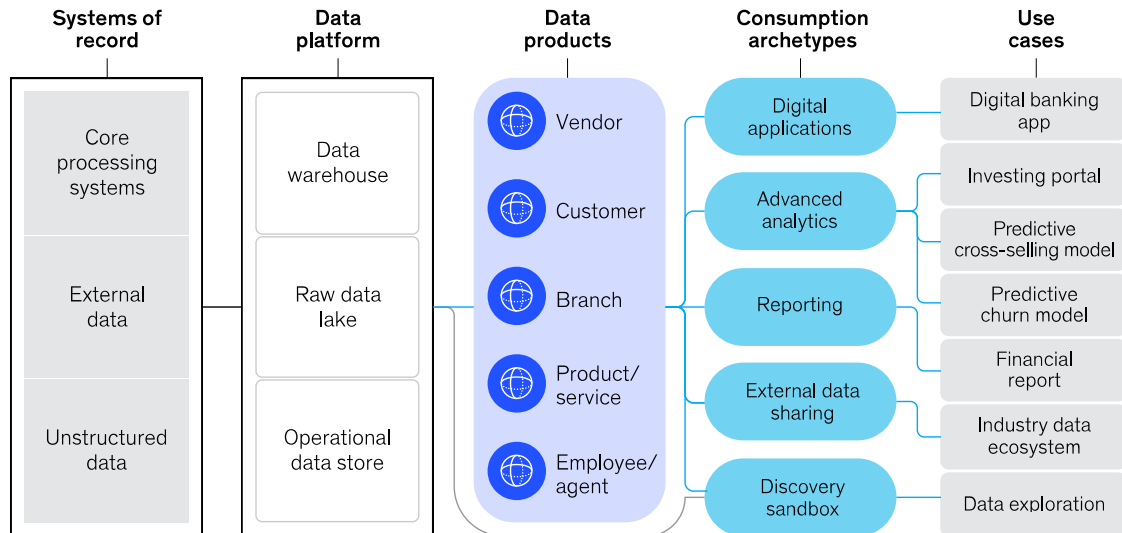
Source: "A better way to put your data to work," *Harvard Business Review*, June 2022
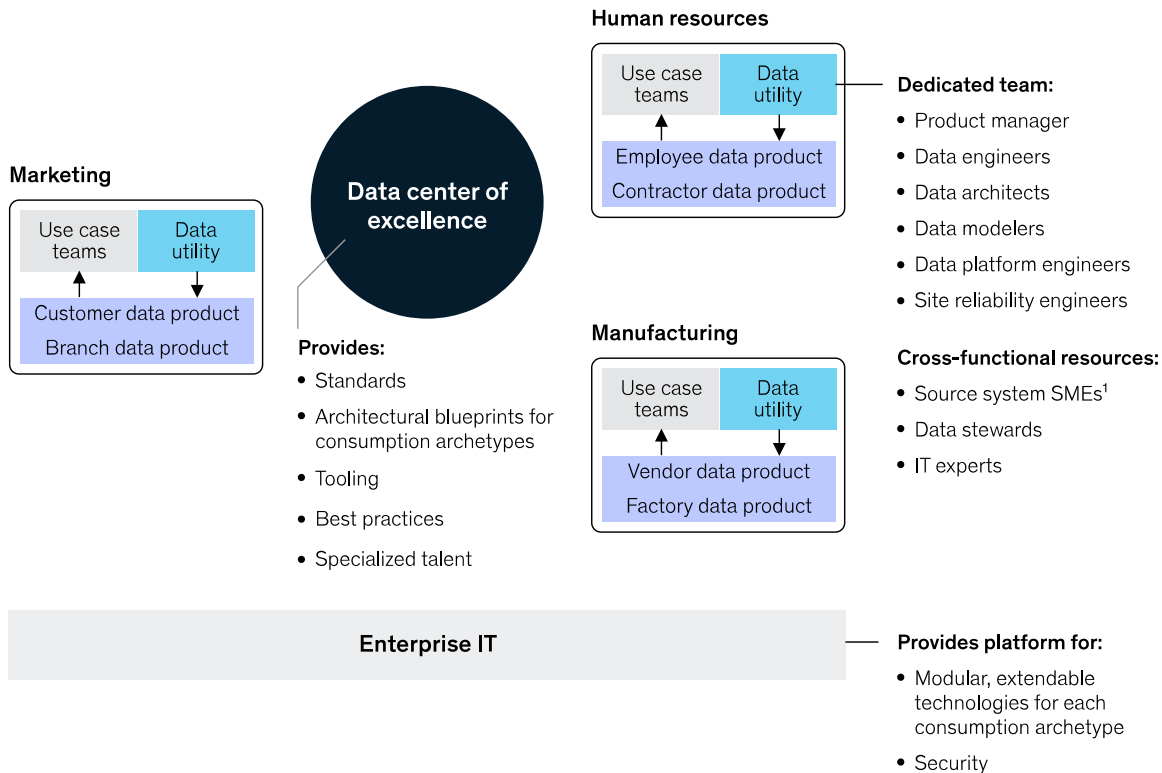
McKinsey
& Company

The benefits of this approach can be significant:

- New business use cases can be delivered as much as 90 percent faster.

- Total cost of ownership, including technology, development, and maintenance costs, can decline by 30 percent.

- The risk and data-governance burden can be reduced.

Getting started with data products

Success in product development requires an operating model that ensures dedicated management and funding, the establishment of standards and best practices, performance tracking, and quality assurance. Success with data products is no different.

Managing data like a product requires the right operating model.

Human resources

Use case teams | Data utility
Employee data product
Contractor data product

Dedicated team:
- Product manager
- Data engineers
- Data architects
- Data modelers
- Data platform engineers
- Site reliability engineers

Cross-functional resources:
- Source system SMEs[1]
- Data stewards
- IT experts

Marketing

Use case teams | Data utility
Customer data product
Branch data product

Data center of excellence

Provides:
- Standards
- Architectural blueprints for consumption archetypes
- Tooling
- Best practices
- Specialized talent

Manufacturing

Use case teams | Data utility
Vendor data product
Factory data product

Enterprise IT

Provides platform for:
- Modular, extendable technologies for each consumption archetype
- Security

[1] Small and medium-size enterprises.

McKinsey & Company

- *Dedicated management and funding.* Each data product should have a product manager and a team consisting of data engineers, data architects, data modelers, data platform engineers, and site reliability engineers who are funded to build and continually improve their product and enable new use cases. These teams should sit within a data utility group inside business units. This organizational structure gives them ready access to the experts they need (including business subject-matter, operational, process, legal, and risk experts) to develop useful and compliant data products. In addition, it gives the teams access to user feedback, which helps them continue to improve products and identify new uses.

- *Standards and best practices.* We find organizations are most successful when they institute standards and best practices for building data products across the organization. This work is typically handled by a data center of excellence. Establishing standards and best practices includes defining how teams will document data provenance, audit data use, and measure data quality, as well as designing how the necessary technologies should fit together for each consumption archetype so they can be reused across all data products.

- *Performance tracking.* To confirm that their products meet end-user needs and are continually improving, data product teams should measure the value of their work. Relevant metrics may include the number of monthly users for a given product, the number of times a product is reused across the business, satisfaction scores from surveys of data users, and the return on investment of use cases enabled.

- *Quality assurance.* Because quality issues can erode end-user trust and retention, data product teams closely manage data definitions (for instance, whether the definition of customer data is limited to active customers or includes active and former customers), availability, and access controls that meet the right level of governance for each use case. To confirm data integrity, they work closely with data stewards who own data source systems.

For a deeper look at how leaders can manage data as they manage a product, read "A better way to put your data to work ," on hbr.org *.*

ABOUT THE AUTHOR(S)

Veeral Desai is a senior expert in McKinsey's Sydney office, where **Tim Fountaine** is a senior partner; **Kayvaun Rowshankish** is a partner in the New York office.

<div style="text-align:center">

Talk to us

</div>

EXPLORE A CAREER WITH US

<div style="text-align:center">

Search Openings

</div>