



# Beyond platforms and cloud: Modernizing data for rapid value creation

**/thoughtworks**

Design. Engineering. AI.

<b>Introduction</b>	<b>3</b>
<b>Where modernization efforts go wrong</b>	<b>4</b>
<b>The answer: A new, product-oriented approach to data modernization</b>	<b>6</b>
<b>What is a data product?</b>	<b>6</b>
<b>Rapid value creation in action</b>	<b>8</b>
<b>Holistic evolution requires holistic modernization</b>	<b>11</b>
<b>Authors</b>	<b>12</b>

## **Everyone's investing in data modernization. But few are seeing amazing results**

A 2024 study by [MIT Research](#) found that over 54% of the companies surveyed had either completed a major data modernization initiative or were in the process of data modernization. Another 23% were planning to initiate data modernization within the next two years.

As you might guess, a lot of that investment has been driven by growing interest in AI. Today, everyone wants to get the most out of AI, and that demands a robust, modern data foundation. In fact, modern data foundations are an essential element for any organization that wants to scale out AI from POC to production and start realizing its value at scale.

The trouble is, many of those investments aren't delivering positive ROI. That's a big problem — not only because businesses need those returns to survive, but also because it means thousands of organizations aren't achieving their data goals.

They've funnelled resources into data modernization, but still can't reliably support autonomous, real-time decision-making, or turn their data into value easily. It's the worst of both worlds.

In this whitepaper, we'll unpack where data modernization efforts are going wrong, and explore how you can get things back on track.

## Where modernization efforts go wrong

Data modernization isn't easy. The journey can be long, and there are a lot of mistakes organizations can make along the way. Here are the three we most commonly see:

### Taking a “build it and they will come” approach

One of the common failure modes for traditional enterprise data projects we've seen has been adopting a “build it and they will come” approach. The approach is based on the hypothesis that if you take the time to create a complete and cleansed repository of all enterprise data, you'll be able to satisfy business stakeholders and deliver analytical solutions at the speed they require. It might take more time initially, but that upfront investment will pay off in the long run.

The problem is that the process of collecting, structuring and cleaning data takes much longer than anyone expects. These programs drag out for months, or even years, during which the users who need the data have to find other sources and solutions. Eventually, the business loses patience with the data team, redirects investment, and often restructures the data organization.

Instead, Thoughtworks takes and advocates for an incremental, use-case driven approach. By showing incremental progress and ROI early, we keep critical stakeholders engaged and supportive of our efforts throughout long data modernization journeys.



## **Not making data quality and interoperability a priority for everyone**

One of the key challenges almost all organizations face during and following data modernization is that they don't put enough focus on creating high-quality, usable data assets from their operational data. The domains closest to data assets — the people using them every day — aren't driven to make their data consumable by other teams and systems.

So, as you might expect, data silos persist. Accessibility and interoperability remain restricted across the organization and powerful cross-domain AI use cases can't be adopted with confidence, because data isn't standardized throughout the business.

## **Failure to modernize engineering practices**

Adopting new capabilities in the cloud won't help you achieve much if your engineering teams don't evolve their practices to incorporate those capabilities. New tools should be paired with new ways of working, creating opportunities to streamline tasks and empower everyone to get the most from modernized technology.

Without this kind of change, lots of teams simply end up replicating the inefficiency of their current operations and processes in the cloud. That means they don't realize any measurable value from their investments, so their modernization efforts grind to a halt.

Ideally, before selecting your technology, you should start with a view of exactly how you'd like engineering practices to evolve. It's best to choose platforms and capabilities that can bring your desired processes and outcomes to life, then ensure teams are supported through that evolution and driven to change their behaviors and practices.

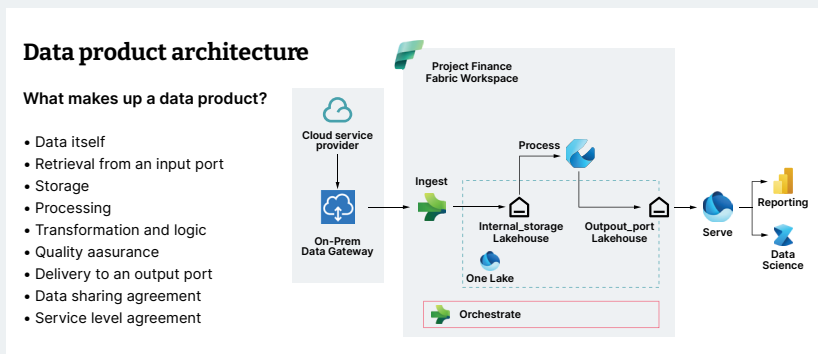
## The answer: A new, product-oriented approach to data modernization

One thing is clear: traditional approaches to data modernization are missing the mark. What organizations need now is a holistic approach that brings together the technological, architectural, process and cultural changes required to enable successful data modernization.






Back in 2020, Thoughtworkers first introduced a new architectural paradigm called Data Mesh. The principles behind it are all aimed at helping organizations create a future-ready data architecture that makes it easy for everyone in the business to access, use and generate maximum value from their data — especially the principle of Data as a Product (DaaP).

## What is a data product?

Data products are processed, analyzed and curated data sets that provide solutions or insights and are packaged as self-contained units with their own independent lifecycle. Typically, they contain both the data required to meet a specific use case and any supporting resources (such as models or capabilities) needed to execute that use case.



Managing data as products that are owned and maintained by the domains closest can help you:

-  **Drive data quality at scale** by making it a priority and responsibility for everyone, and incentivizing teams to ensure their data is clean, valuable, and ready for use
-  **Standardize data across the business** by using platform-level guardrails to ensure data products are interoperable and consistent by design
-  **Make data more visible and discoverable** so that whenever a team needs to use it, they can help themselves to it and get started right away
-  **Break down silos** and enable ease of accessibility through self-service data platforms
-  **Embed new processes and engineering practices** and align everyone around a product-oriented way of managing and working with data





DaaP helps solve multiple challenges that lead to ineffective data modernization efforts. It's also aligned with the goals of modern businesses and their data leaders, including:

-  **Getting data AI-ready** by structuring it in well-maintained, high-quality products that are ready to be used in AI use cases
-  **Enabling cross-functional innovation** and breaking down the silos between teams so they can benefit from each other's data and insights
-  **Supporting rapid value creation** by reducing the time it takes to make data actionable, and making data more accessible so more teams can use it to generate value

However, the DaaP concept alone isn't enough to ensure successful and impactful data modernization.

## Building around (and beyond) DaaP

Much like traditional approaches to data modernization, the value of DaaP hinges on what you build around it. It must be supported by:

-  **Robust self-service data platforms** that make data products visible and accessible across the business
-  **AI-powered capabilities** that accelerate routine data management tasks and value creation
-  **A shift left in testing strategies** that introduces tests earlier in the development cycle to help identify the most valuable data products and move on quickly if products lack impact
-  **Continuous innovation, integration and delivery** so all three become always-on business processes, creating a perpetual value engine that continuously extends the value the business can get from its data

## Rapid value creation in action

By applying modern data engineering, AI, and platform capabilities alongside the principles of Data as a Product, we've helped numerous organizations break down data silos, improve governance, and accelerate insights. They've achieved some remarkable results, including up to 4x faster decision-making, 20-30% lower platform costs, and rapid scaling of AI-driven solutions.




Here's a snapshot of some of the impact we've seen this approach deliver.

### **Gilead Sciences: Building a platform foundation for AI-driven innovation**

At Gilead, we supported the design and implementation of a new enterprise-wide data platform called Gilead DnA. The platform

was built to be highly scalable and adapt to Gilead's evolving data demands and use cases. But more importantly, it gave the organization's engineers and research teams a secure way to process data themselves and 'talk to their data' using built-in AI capabilities.

This platform provided the foundation for numerous innovations, including digital twins for research and AI-driven capabilities that have reshaped the company's biologics labs and processes, ultimately helping Gilead:

-  Accelerate decision-making and research with real-time data analytics and access
-  Increase productivity by digitizing a wide range of paper-based processes and boosting data accessibility
-  Optimize costs through smarter resource allocation and continuous improvement

## **Roche: Building a flexible, future-ready data architecture**

Thoughtworks helped Roche undergo a complete data transformation, building and implementing its own data mesh architecture. As part of the journey, we helped Roche's central teams and domains embrace Data as a Product, and build the platforms and capabilities to manage their data as domain-oriented products.

Crucially, this technical evolution was supported by widespread process and cultural change. New engineering practices were embedded at every level to support the maintenance and expansion of the data mesh. Domains were given defined responsibilities for the governance of their data products, helping increase data quality at scale. And the architecture was agile by design, so it can evolve alongside Roche's shifting needs and enable continuous value creation as new data-driven opportunities emerge.

## **Bayer: Transforming how preclinical data is managed and accessed**

Bayer manages huge volumes of data and reports on preclinical trials. In the past, much of that data was spread across siloed systems and teams, making it difficult to find and leading to a lot of testing and effort duplication.

To solve that challenge, we helped Bayer build and implement PRINCE (Preclinical Information Center), a new platform that acts as a one-stop shop for all preclinical data, uniting previously siloed data sources.

PRINCE's primary data product contains the results from thousands of toxicology studies alongside in vitro and in vivo bioassay data and compound information. Users have access to structured and unstructured information, whether they're accessing the historical database through APIs, dashboards, or the platform's search interface — enabling researchers to inform their work with what has been done and learned before.

Embedded agile engineering practices ensure the platform is continuously evolving and able to support a growing number of use cases for preclinical data. Since its launch, we've also introduced a generative AI-powered chatbot that's transformed how researchers access and interpret vast volumes of structured and unstructured preclinical data.

**“PRINCE represents a major milestone in the digitalizing of our preclinical data domain. We think that it can serve as a blueprint for a future domain-centric decentralized data landscape in R&D.”**

**Jonas Münch**

**Head of IT for Safety & Pharmacology at Bayer**

## Holistic evolution requires holistic modernization

Data is the key to efficiency, innovation and value creation. If you want to get the most from your data at scale, you need a modern data architecture, supported by modern engineering practices and proven strategies for improving data quality and governance.

By applying Data as a Product principles, modern engineering practices, and new data platforms and architectures, Thoughtworks enables holistic data modernization. We look beyond the technical elements of your data environment to consider the structural, domain-level and individual-level changes that must happen to enable rapid value creation from data across your business.

Through our unique modernization approach, we can help you:

**Reimagine how your teams work with data** and contribute to better data quality through more robust data capture and structuring

**Break down long-standing data silos** and bring standardization to enterprise data environments, so everyone can make better use of each other's valuable data

**Become AI-ready** and build valuable data products that can easily be applied to enable new AI and ML use cases

**Harness AI for data management** and use it to support human data access and management

**Enable continuous value creation from data** and turn your data organization into an always-on value creation engine by shortening the loop from ingestion to action

If you're looking for new ways to unlock the full value of your data, visit our [website](#) or reach out. We'll help you explore your needs, and identify which strategies, practices and approaches may be best suited to help you achieve your data goals and lay a robust foundation for years of data and AI-driven innovation.

## Author



**John Spens**

Director, Data and AI - North America



We are a global technology consultancy that delivers extraordinary impact by blending design, engineering and AI expertise.

For over 30 years, our culture of innovation and technological excellence has helped clients strengthen their enterprise systems, scale with agility and create seamless digital experiences.

We're dedicated to solving our clients' most critical challenges, combining AI and human ingenuity to turn their ambitious ideas into reality.

[thoughtworks.com](https://www.thoughtworks.com)