

# The Definitive Guide to Industrial DataOps



# Contents

4

Introduction:  
**The Time for Industrial  
DataOps Is Now**

16

Chapter 1:  
**Your Digital Maturity:  
Are You Ready for  
Industrial DataOps?**

104

Chapter 4:  
**Industrial DataOps  
in Action**

132

Conclusion:  
**Industrial DataOps Is  
the Future**

44

Chapter 2:  
**DataOps as a  
Discipline: Extracting  
Value from Data**

142

Appendix:  
**Industrial DataOps  
RFP Guideline**

72

Chapter 3:  
**Principles of  
Industrial DataOps**

157

Cognite:  
**About the Authors**

---

## Introduction

# The Time for Industrial DataOps Is Now



# The Mother of [Re]Invention

The year 2020 forced certain industrial sectors, once seemingly invincible, to grapple with their mortality. The economy was shattered. Oil prices fell below zero for the first time in history. Airline stocks bottomed out. Millions of jobs were lost.

Worst case scenarios were no longer just hypotheticals: they would be distinct possibilities without massive efforts to transform, to address vulnerabilities, build resilience, and further digitalize. In fact, had the pandemic happened just 10 or 20 years earlier, before the advent of the technology that saved lives, jobs, and markets, we could have seen far worse.

Suddenly, once-analog traditional industries like oil and gas, manufacturing, power and utilities, began adopting digital tools at a faster pace.

Covid-19 forced most workers home, so we tested remote work en masse—and largely with success. We also piloted remote operations and learned more about honing technology to create real value out of industrial data. Necessity was the mother of a wide-scale reinvention.

Meanwhile, these same industries started to address public and financial pressure to own up to environmental impact and unchecked emissions. After all, unprecedented climate-related disasters were playing out before our eyes, adding insult to the pandemic's injury, holding our gaze on frightening truths. Australia, California, and Colorado burned. Power grids were compromised. A record number of hurricanes hit the US Gulf Coast. Southeast Asia dealt with record floods. Siberia melted.

## Eyes Wide Open

However jarring these realizations, opportunities come from having our eyes wide open. The collective progress that industrial companies made in 2020 came largely out of necessity and a real sense of urgency, but were nonetheless impressive. By and large, digital transformation is no longer seen as an overused PR term, nor a risky CAPEX move.

**Adopting industrial software, digital tools, robotics, and new agile ways of work are gaining ground. These are areas where asset-intensive industries had been notoriously behind the curve until now.**

The value of industrial data, liberated and contextualized by a new generation of software solutions, is now widely recognized. And finally, a willingness to share, open and digest such data is catching on across our industries. The time is now to capitalize on this.

The foundation for this necessary, sustainable reinvention is here now. It is built of public and financial incentives, competitive motivation, survival instincts, innovation, and genuine goodwill. Required now, and at scale, are the know-how, the technology, and the tools to do the actual transformation work.



# Industrial DataOps: A New Tool in Our Box

---

One absolutely essential tool is the new, rapidly growing discipline of DataOps, and more specifically, the novel approach of **Industrial DataOps**.

Just a twinkle in the eye of industry today, we expect that this will become one of the single most important vehicles turning industrial data into tangible value. In doing so, Industrial DataOps will become a driving force in industrial transformation.

Because Industrial DataOps is a truly innovative and transformative new approach, it takes the rest of this six-chapter book to fully unpack what it is, why it's important to your company and to your industry, why we need it now, and how to begin implementing it. Your 101 into Industrial DataOps starts here.

“

**Industrial DataOps in 2020 is as new and unknown as DevOps was in the 2010s, as business intelligence was in the 1990s, and as human resources was in the 1890s.**

# What Industrial DataOps Is

The core principles of Industrial DataOps will be developed in more detail in the following chapters. So here we'll focus on simply providing a basic outline. Let's begin by defining what DataOps is.

DataOps can be defined as a new discipline which typically includes a team of data experts (think: data scientists, analysts, architects, and the like) who exist to "provide the tools, process and organizational structures to support the data-focused enterprise."<sup>1</sup>

DataOps teams are tasked with bringing data to data customers throughout the organization, internal and external, and making their data do more for the company.

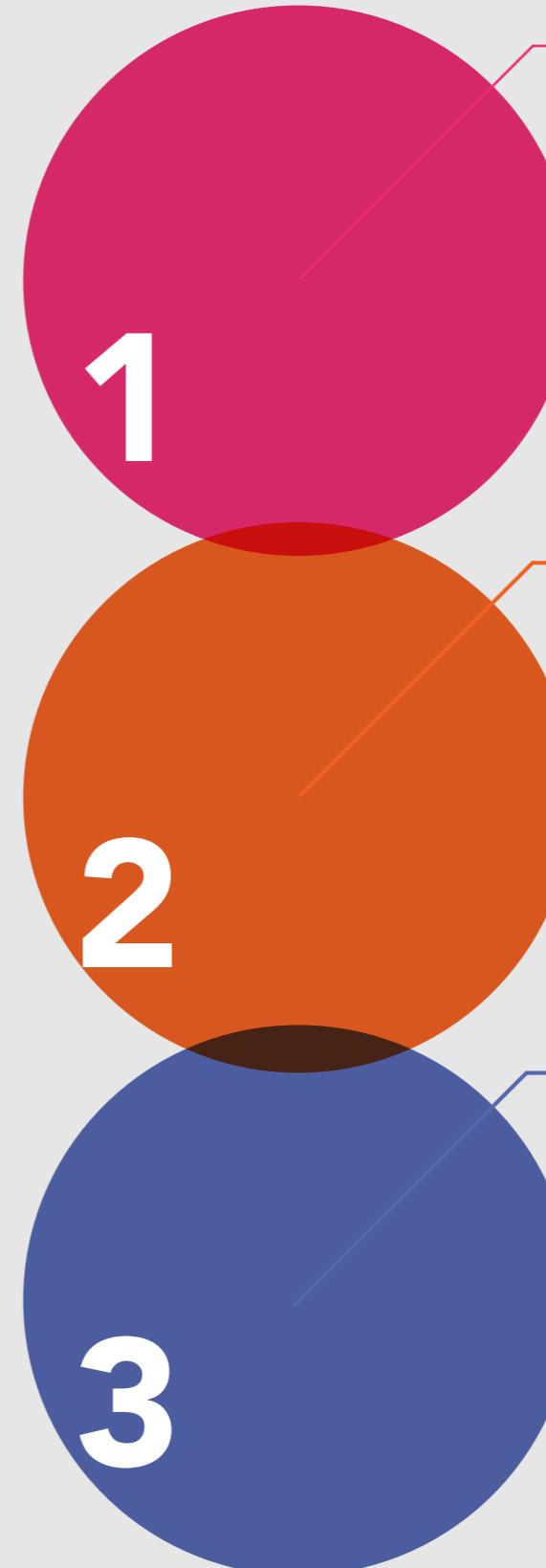
According to Gartner, "DataOps is a collaborative data management practice focused on improving the communication, integration and automation of data flows between data managers and data consumers across an organization."<sup>2</sup>

For Forrester, "DataOps is the ability to enable solutions, develop data products, and activate data for business value across all technology tiers from infrastructure to experience."<sup>3</sup>

Two strong common threads across these definitions are the importance of collaboration, and the focus on integrating data across different parts of an organization.

**Industrial DataOps is about breaking down silos and optimizing the broad availability and usability of industrial data. By this we mean the data generated in heavy asset industries including oil and gas, power and utilities, manufacturing, and shipping.**

# 3 Key Truths About Industrial DataOps



## **Industrial DataOps depends on collaboration with domain experts.**

Individuals and interactions (far more than processes and tools) are essential to make data valuable and useful for the data consumers across an organization: the domain experts in different fields and departments. It's important to remember that DataOps is a practice, a way of engaging and collaborating across the organization to both share and reap greater value from the data.

## **Data is only as valuable as the analytics behind it and the scale of people who use it.**

The convergence of data and analytics have made Industrial DataOps an operational necessity. But the data requires context if it is to be used broadly by non-data doctors. Only by automating the data process and creating one central, contextualized source of truth, can we ensure the live data triumphs over its predecessors (the static documentation and reports of yesteryear) in the decision-making process.

## **Extracting the value of the data requires an agile approach.**

Industrial DataOps isn't about documenting, reporting, or extensive up-front design. It's a far more agile process in which experimentation, iteration, and feedback are essential. Creating business value isn't a one-way transaction between data scientist and department. It's a joint effort that will require both parties to participate, share, and develop solutions that hold transformative potential. Just as the data is alive, so are the means of working with it.

# What Industrial DataOps Is Not

Industrial DataOps is not DevOps. DevOps has been around for some time and, despite some superficial commonalities, DevOps is very different. Both are methodologies used to enhance operational practices, but that's where the similarity ends.

The focus for DataOps is the delivery of business-ready, trusted, actionable, high-quality data, available to all data consumers or domain experts throughout the organization. One goal is automation efforts, centered on data governance and integration.

Another goal is alignment between IT system support, operations, and the business. The focus for DevOps, on the other hand, is software and application development. Automation efforts center on the development cycle, software delivery processes, and waste elimination. The alignment of developers, operations, and the business is a major aim of DevOps.



## Why Industrial DataOps and Why Now?

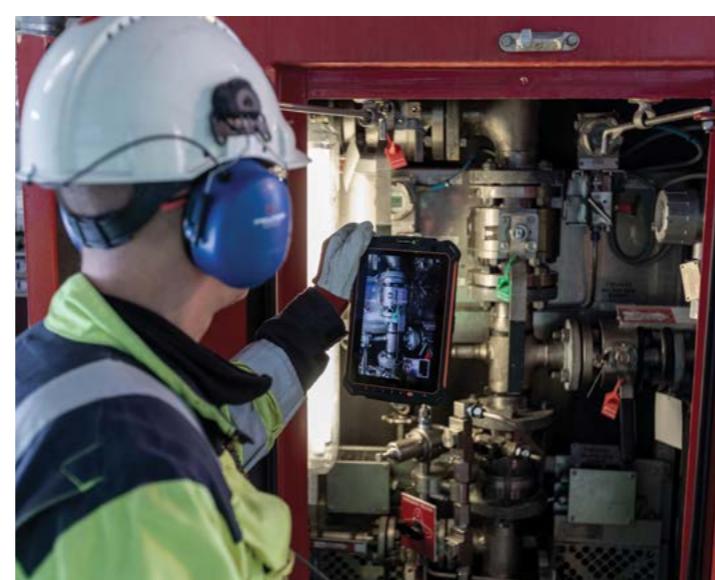
Industrial DataOps is fundamental to driving up data literacy: the ability to read, write, and communicate data. Data literacy, in turn, is key to obtaining optimal value from data and digital. Therefore, the more organizations embrace Industrial DataOps, the better equipped they will be to truly harness the transformative potential of data, getting us one step closer to our shared end goal of sustainable, efficient, safe industrial operations.

The above has been a starting point for exploring how data can enable asset intensive industries to develop solutions and data products, and to extract value across an entire business. It's almost impossible to imagine that not long ago, data was a protected commodity, something to be kept to oneself or even used as a bargaining chip. This was, thankfully, a short-lived narrow-mindedness, brushed aside by visionaries who saw the amazing potential of data. Their vision of the future has spread to the rest of us, sparking the emergence of Industrial DataOps as our lifeboat in a vast sea of data.

So why now? Because delaying the deployment of advanced data operations and new ways of working will only delay the transformation of your company, and by extension, your industry, making achieving your end goals slower, harder, and more expensive.

**More and more heavy-asset industry companies are waking up to the power of data and its ability to inform decision-making, transform operations, and enhance sustainability.**

As this happens, we hope that you come to see this as an essential guide to rolling out Industrial DataOps in your organization.



---

## Chapter 1

Your Digital Maturity:  
**Are You Ready  
for Industrial  
DataOps?**



# Digital Maturity Is a Key Metric of Digital Success

Any organization that wants to adopt and enjoy the benefits of Industrial DataOps needs to consider the state of its digital maturity. It's clear that outsized ROI from digital investments is more likely to come out of sustained innovation and long-term digital strategy, rather than just one or two quarters of focused effort.

This doesn't mean that deploying more use cases isn't important and that short term-value isn't possible, but rather that digital programs and operational momentum are difficult to get started and can fail quickly for a myriad of reasons.

**Outsized ROI from digital investments is more likely to come out of sustained innovation and long-term digital strategy, rather than just one or two quarters of focused effort.**

This is exactly why so many organizations are frustrated right now; they expected outsized short-term payoffs, not fully realizing the longer-term play. To be fair, the broader digital transformation market did overpromise the immediacy of value from digital, but the solutions market remains accurate regarding the significant potential of advanced digital technologies. This is a classic case of "when," not "if."

Symptomatically, this disconnect manifests itself in terms of stalled digital transformations, as the evidence from McKinsey (Fig. 1) demonstrates compellingly.

Across industry, stalled progress happens most often during the scaling phase, with 62% of the root cause being factors that are within the organization's near to medium-term control. Developing an operational proof of concept (POC) for example, has become much simpler in the last two to three years, while scaling still remains a challenge.

This points to a very important parallel, however: POCs are short-term projects while scaling is long(er)-term. What's missing in this paradigm today is the lack of a proper people, process, and platform flywheel to carry and build digital momentum forward from phase to phase.



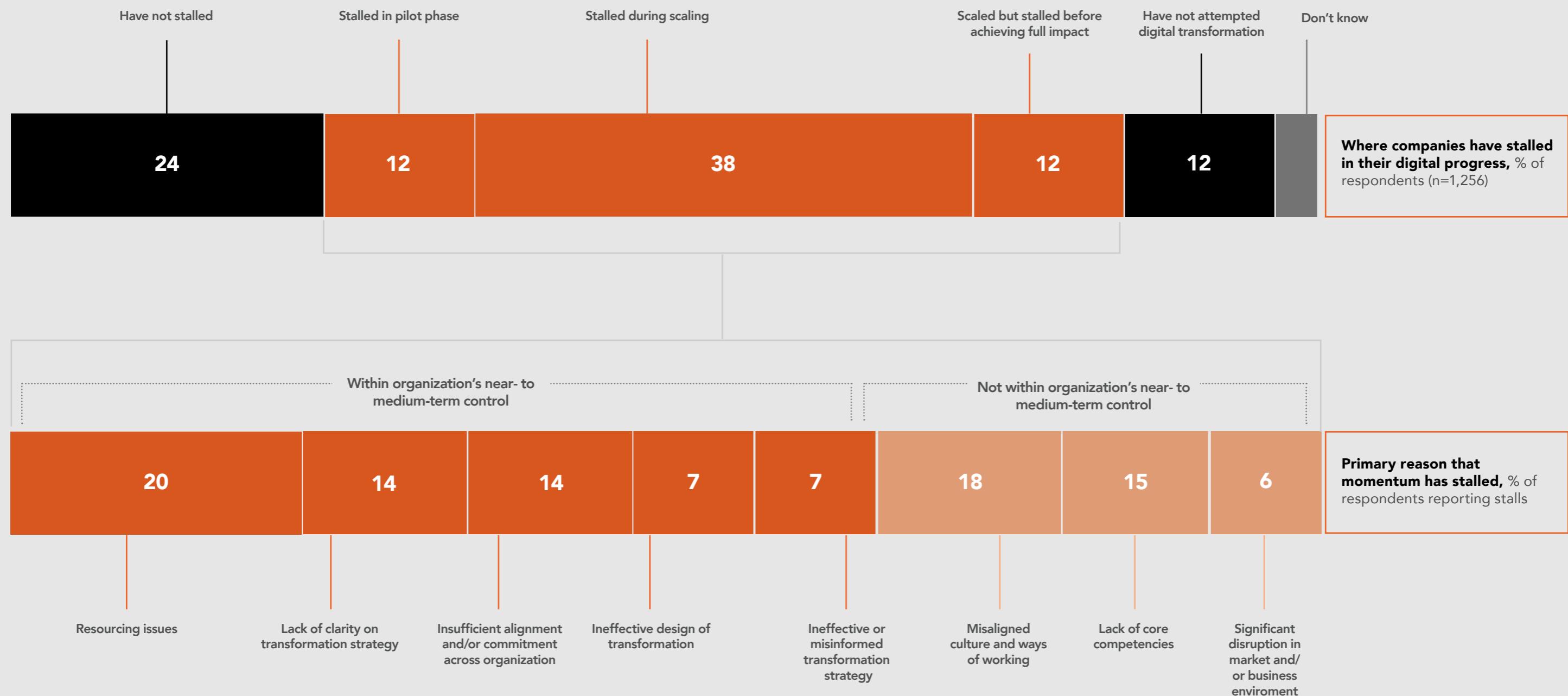
**The measures and KPIs used in decision-making today are so entrenched in showing short-term value that they neglect to incentivize the next phases of digital transformation that support scaling and beyond.**

Here, new long-term metrics around operational adoption, cost of analytics, and digital spend as a percentage of operations become a much more important part of the equation. They can be bundled into a new and arguably more critical measure: digital maturity.

Digital maturity, when codified as a practice and quantified as a metric, represents a sustainable way of measuring progress in aggregate, and a better measure of overall industrial digital transformation health.

## Figure 1: Stalling in the Digital Process

Most respondents say their digital progress stalled, with the majority of stalls caused by factors within their organization's control.



# Defining Digital Maturity

---

The first shift that must occur involves realizing that the digital maturity curve is **exponential**, not linear. In broad terms, early maturity consists of many linked or unlinked data “pieces” or building blocks that, when combined with high-cost internal or third-party services, start to create silos of opportunity and kick-start small-scale projects and POCs. As the organization matures, internal or vendor-delivered application frameworks start to create some repeatability, especially when combined with a data analytics platform.

This platform serves as a bridge to high maturity, because it carries forward the momentum and infrastructure to develop data analytics catalogs and libraries that can then be deployed with fewer services and at lower marginal costs. It is during this transition that outsized ROI starts to develop.



**Figure 2: The Relationship Between Digital Maturity and ROI**

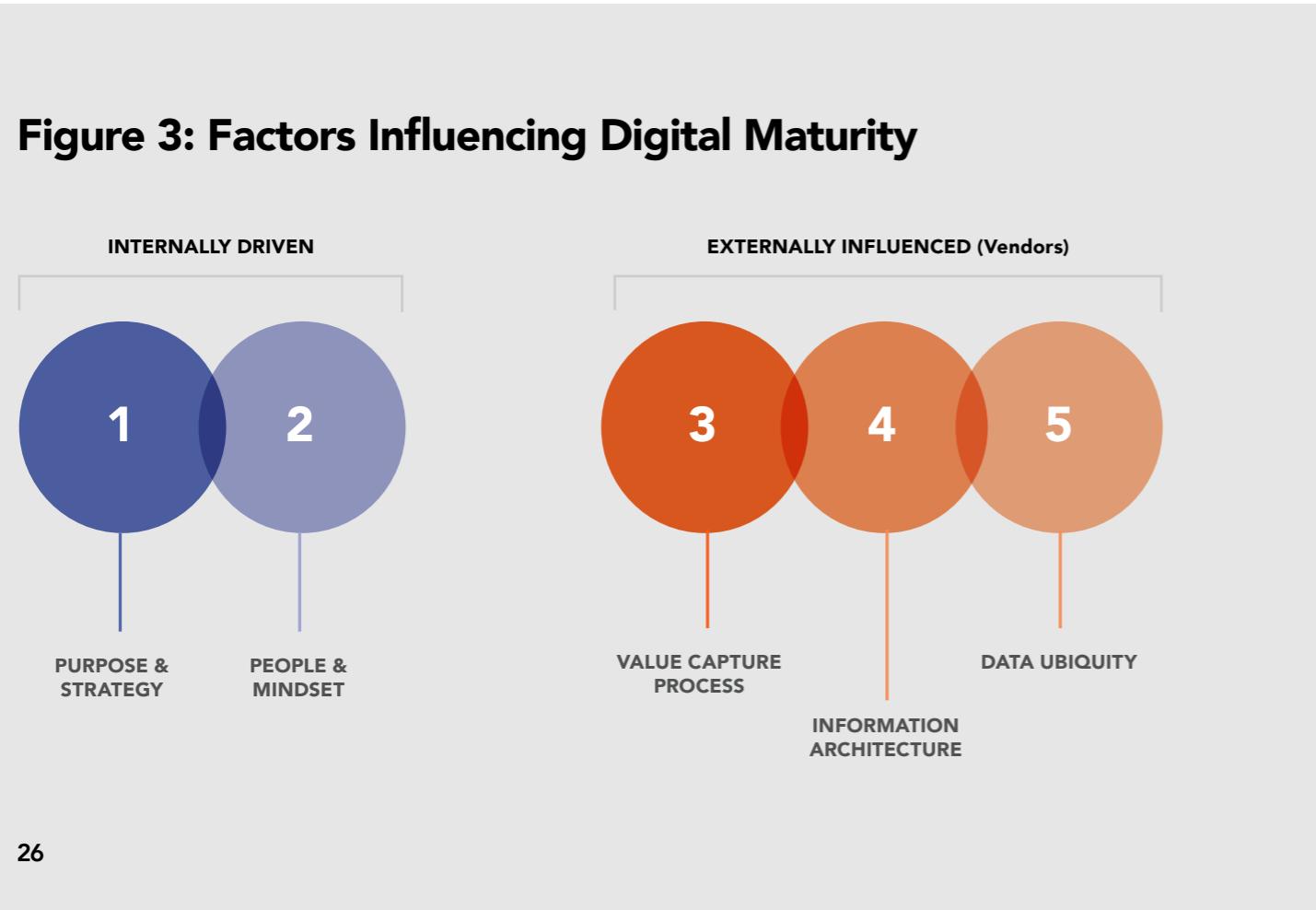


# What Does High Digital Maturity Look Like?

While building organizations in support of digital maturity can look different at a company level or within an individual line of business,

high-maturity organizations share some common attributes, segmented by internal versus external factors (Fig. 3).

**Figure 3: Factors Influencing Digital Maturity**



INTERNALLY DRIVEN FACTORS		
1 PURPOSE AND STRATEGY	Digital ambition is clear, defined, and strategic. Digital goals are time-bound for momentum and velocity.	IT/OT are equally accountable for digital success. Digital strategy and KPIs are communicated frequently.
2 PEOPLE AND MINDSET	The digital program is intentional, structured, and visible. Digital leadership is empowered and accountable.	Digital capabilities are progressive and diversified. Digital culture encourages and incentivizes citizen contribution from domain experts and non-data specialists.
EXTERNALLY DRIVEN FACTORS		
3 VALUE CAPTURE PROCESS	Value pools are strategic, well-defined, and measured. New and existing use cases move fluidly from concept to scale.	Delivery combines established processes with hyper-automation. Value is captured throughout projects and communicated to the organization.
4 INFORMATION ARCHITECTURE	Digital ecosystem is highly structured and cohesive. Data is available securely on demand with high quality and access control.	Operational data is highly utilized with site- and corporate-level granularity. Broad portfolio of low-code tools exist for ad hoc or specialist use.
5 DATA UBIQUITY	80%+ of structured and unstructured (IT/OT/ET) is highly utilized. Data standardization in source and format is a priority across sites and assets.	Data quality and integrity checks are automated and perform with high precision. Data discovery is simple, intuitive, and approachable for non-SMEs.



## **When an individual organization aspires for and invests in achieving digital maturity as a strategy, organizational performance improves.**

Recent research by Deloitte clearly associated more digitally mature organizations with greater efficiency, revenue growth and net profits above the industry average, as well as higher product/service quality, better customer satisfaction, and stronger employee engagement.<sup>4</sup>

Digital maturity therefore paves the organization's way for success, but this potential may not be realized if the digital competencies of one department are significantly more evolved than those of the rest of the company.

Contrast this with a company-wide mandate to invest in digital maturity, and the resulting momentum shifts from focusing on the nuance and noise of every digital project—whether successful or not—to addressing root cause.

# Digital Traps to Avoid

Now let's unpack a few examples of the traps organizations can fall into on the quest for digital maturity.

## Digital Trap 1: Skunkworks Projects

Even in low-maturity organizations, business problems or opportunities will eventually get addressed. This typically happens in an ad hoc manner either once they become frustrating enough or it becomes impossible to ignore the potential perceived value. Here, opportunistic employees are likely to take the initiative and try to solve the problem given their existing skills and tools.

Take for example, an operator who is frustrated because they can't easily compare historical events with the time series data for a particular asset. If they had access to the asset's history and a record of time-stamped events, they could create a spreadsheet-based report to link the data and visualize it in a simple table.

Many of the problems being solved through these skunkworks (informal innovation) projects are highly valuable. But unless the organization's digital maturity supports an intentional, low-friction path to productionalize the application, only a fraction of value might be realized because the report stays within its individual silo. This points to opportunities to improve digital processes and platforms.

If the question is "What's needed to transform these vanity projects into long-term value?", the answer is once again a higher maturity organization. In this case, the opportunity exists around leadership and process. Leadership must be able to balance needs for signaling with a tight, methodical value assessment and capture process that drives investment and progress on specific operational goals and KPIs.

## Digital Trap 2: Vanity Projects

On another side of the spectrum, the organization may get involved in projects that seem innovative and groundbreaking but never materialize into tangible, traceable business value. Either the project costs significantly more than the value delivered or it introduces a new technology that creates additional downstream issues in the workflow, thus negating any significant benefit.

As new technologies such as drones and robotics are increasingly deployed in operational settings, it's easy to see the challenges of this balancing act. On one hand, the real-world use cases are tangible and valuable, delivering remote access, visibility, and safety to human operators. On the other hand, they serve as clear symbols to customers, markets, investors, and competitors that the company is aggressively pursuing digitalization.

## Digital Trap 3: Elephant Projects

A third circumstance involves projects that are quite sophisticated and solve a real business problem, but require a massive amount of resources and maintenance to sustain over the life cycle. Certain artificial intelligence (AI) and machine learning (ML) projects fall into this category because many digital teams underestimate the effort it takes to fully productionalize some ML-based predictive analytics.

Predictive maintenance on critical industrial assets is certainly a high-impact business problem that is worth solving. To a skilled data scientist, the process for creating initial models and showing predictive results is becoming less complex and more straightforward. But the problems shift as the models get closer to production. Once deployed, they still require maintenance, training, and integrations on an ongoing basis.

Unless the team has already accounted for that long-term overhead, the project costs must come down. Here, high-maturity organizations have been able to build in economies of scale at a data modeling level through investment in platforms, people, and processes. It becomes much less of a burden to carry the remaining overhead from multiple deployed projects, and resources can be shifted towards the next high-value project.



# The Risks of Inadequate Digital Maturity

Over 100 oil and gas vendors went bankrupt in 2020 due to a number of factors such as the price of oil, shrinking demand, and a record surplus of reserves. This is obviously an extreme example with many variables, but it speaks to the risk of not fully leveraging data to identify hidden opportunities, change business models, or adapt to rapidly changing economic scenarios.

Pacific Gas and Electric (PG&E), California's largest utility provider, provides a cautionary tale of what can happen when an organization fails to use data fully to its advantage. In 2018, transmission line faults sparked a deadly, large-scale fire, one of several blamed on the company's equipment.<sup>5</sup> But better data and digital systems could have alerted engineers to the risk, preventing the loss of life and billions of dollars of damages.

Although the organization was known for embracing innovation and spearheading new technologies, PG&E's motivations and digital infrastructure were not necessarily designed to enable these types of operational changes at the required pace. A higher level of digital maturity (and subsequent value from better simulation and condition monitoring) could have helped implement the digital systems needed to mitigate wildfire risks from areas where data is highly available.

At a very practical level, these are the key risks associated with failure to adopt a strategy of increasing digital maturity:

## Overwhelming Complexity:

New data silos are created with each new use case, creating a seemingly never-ending backlog of disparate processes and new technical debt.

## Underwhelmed Board of Directors:

Digital progress becomes increasingly difficult to achieve and quantify, and ROI decreases due to long-term costs.

## Loss of Competitive Advantage:

The digital adoption rate continues to accelerate. Are you at risk of being unseated by a more nimble competitor?

Moreover, these risks compound on each other, causing less-than-ideal strategy and decision-making that further destroy any momentum gained from digital programs.



# The Rewards of Higher Digital Maturity

---

So what are the rewards of pursuing digital maturity? Well, in addition to offsetting the risks above, digitally mature organizations enjoy a host of tangible benefits:

## Top-Line and Bottom-Line Improvement:

The organization can realize market opportunities and new business models while also driving significant efficiencies in day-to-day operations.

## Industry Leadership Recognition:

Not only does this digital leadership boost the brand, but digital programs add no obvious risks.

## Some Insulation From Volatility:

In cyclical or volatile markets, those companies that can use their data in a sophisticated fashion can identify early indicators and get ahead of significant exposure.

## Improved Operational Flexibility:

Data-driven decision-making that equips agility to focus on the highest ROI activity at any given time.

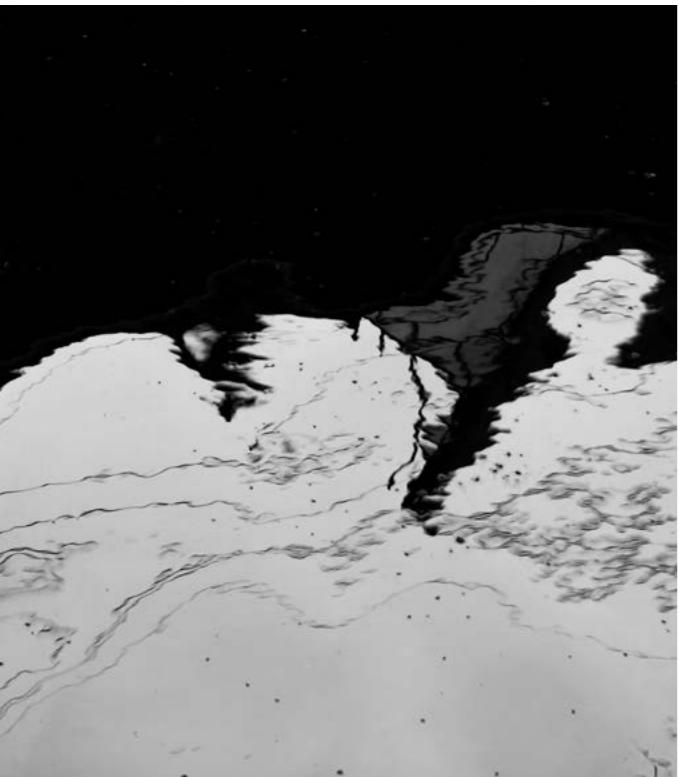
## An Empowered Workforce:

Capable of solving certain problems with confidence that they have the tools and skills to put data to work.

## Lowest Possible Costs of Using Data:

High-maturity organizations tap into economies of scale that increase their ability to innovate while mitigating the risks and costs of failure.

# Organizations Reaping the Rewards of Digital Maturity



There are many examples of industrial organizations embracing the journey to higher digital maturity and reaping the rewards. In the energy sector, British Petroleum (BP) continues to move at lightning speed with a strong digital program<sup>6</sup> reflected from leadership, all the way into each individual business unit.

**BP continues to expand its capacity to make informed strategic bets, fail fast, and quickly iterate on and develop solutions that can be shared across teams.**

**"You need to build your foundation so you can add the technologies that are much more advanced. You need to define the problems you want to solve and then implement those technologies."**

Robert Sentz, Senior Engineering Specialist, 3M

In manufacturing, 3M leads the charge with a very realistic attitude towards digital maturity and summarizes well the challenges in this space. As Robert Sentz, Senior Engineering Specialist at 3M, said, "It's very difficult to just dive right into digital transformation. You must have a solid and robust foundation in place to transform.

"To start the journey, step one is an assessment of what you have today and where you want to go in the future—a road map. Then you need to build your foundation so you can add the technologies that are much more advanced. You need to define the problems you want to solve and then implement those technologies."<sup>7</sup>

A third example comes from power and utilities, a sector known for cautious and late technology adoption due to the high stakes nature of its business. But the New York Power Authority (NYPA) is widely considered to be a model of digital innovation. NYPA has been aggressively pursuing innovation in both IT and operational technology (OT) as a strategy, choosing to risk small failures in service of developing a culture that supports next-generation thinking and technology implementation. The overall mission? To "lead the transition to a carbon-free, economically vibrant New York through customer partnerships, innovative energy solutions, and the responsible supply of affordable, clean, and reliable electricity."<sup>8</sup>



# How Do You Measure?

## Defining a Sustainable Framework

At this point, we must start to convert theory into practice. Let's refer back to the five key factors of digital maturity discussed earlier (Fig. 3). Generated after evaluating a number of organizations in heavy-asset industries and combining the findings with third-party research, these form a sophisticated framework for discussion and evaluation.

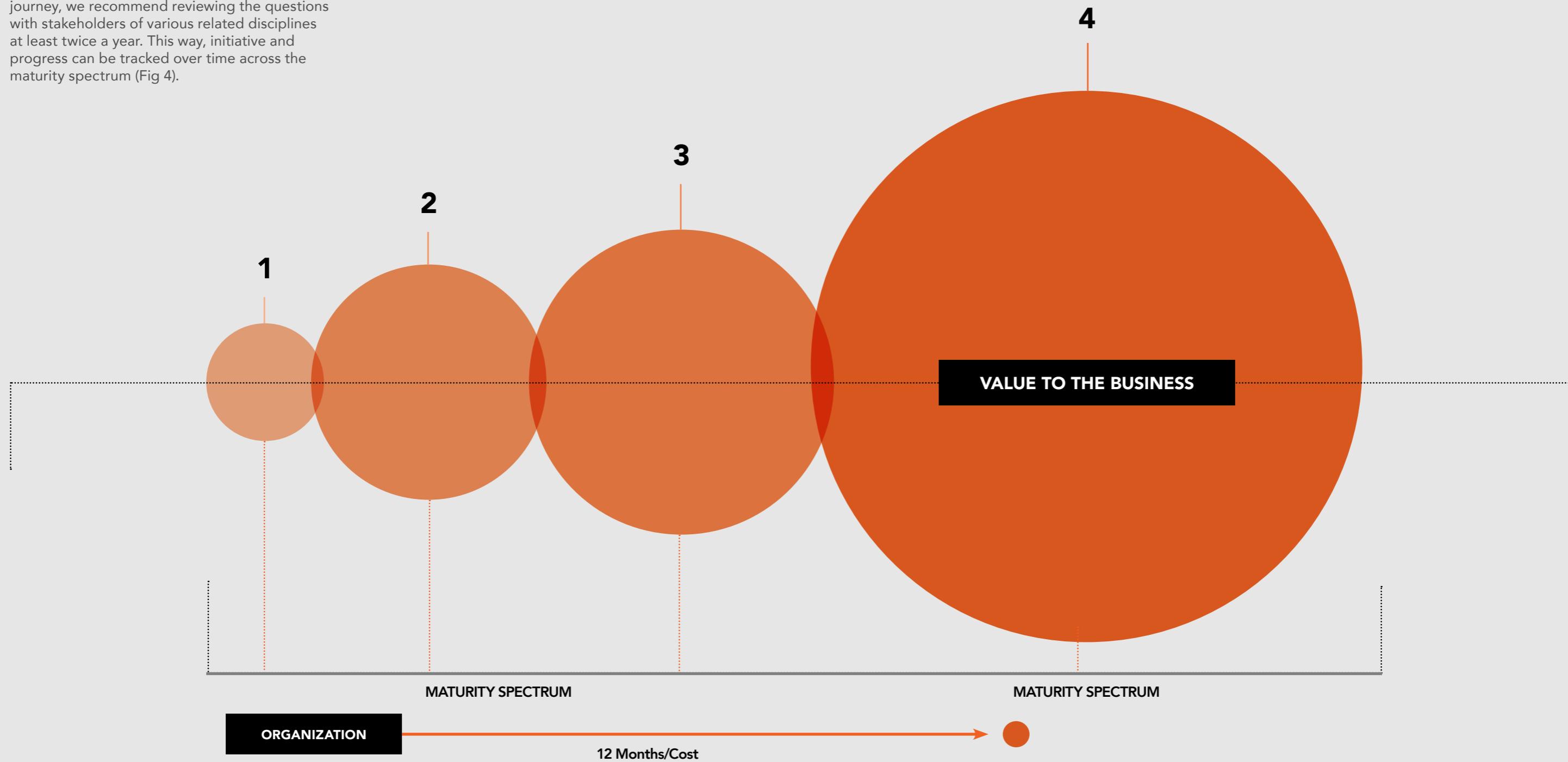
Here, we can further segment the key factors influencing digital maturity into additional parameters that can be measured and weighed together to transform the concept of digital maturity into a diagnostic.

## Digital Maturity Diagnostic Framework

CATEGORY	PARAMETER
PURPOSE & STRATEGY	<b>Digital Ambition: Goals</b> How lofty are the digital goals and the impact on the company's strategic positioning?  <b>Digital Ambition: Velocity</b> How quickly are goals expected to be achieved?
PEOPLE & MINDSET	<b>Leadership</b> How aligned are the different levels of the organization?  <b>Capabilities</b> How advanced/capable is the digital organization: including data scientists, developers, and analysts?
VALUE CAPTURE PROCESS	<b>Value Linked to Strategy</b> How well have areas of value been defined, assessed, and prioritized?  <b>Use Case Trajectory</b> How well/quickly do business problems become solutions?
INFORMATION ARCHITECTURE	<b>Enterprise Architecture</b> How integrated, flexible, and open is the enterprise data stack?  <b>Operational Technology &amp; Data</b> How evolved, structured and managed is the process for maintaining OT source system data?
DATA UBIQUITY	<b>Data Type Diversity</b> What types of data are being used for analytics?  <b>Standardization</b> How well are data standards applied across sites and business units?
	<b>Quality</b> How are data quality issues identified and addressed?  <b>Discovery</b> What is the process for finding data relevant to a problem set?

## Figure 4: Digital Maturity Spectrum

To achieve the best results and to truly have a living, breathing measure of this digital maturity journey, we recommend reviewing the questions with stakeholders of various related disciplines at least twice a year. This way, initiative and progress can be tracked over time across the maturity spectrum (Fig 4).



# Industrial Digital Maturity in Practice

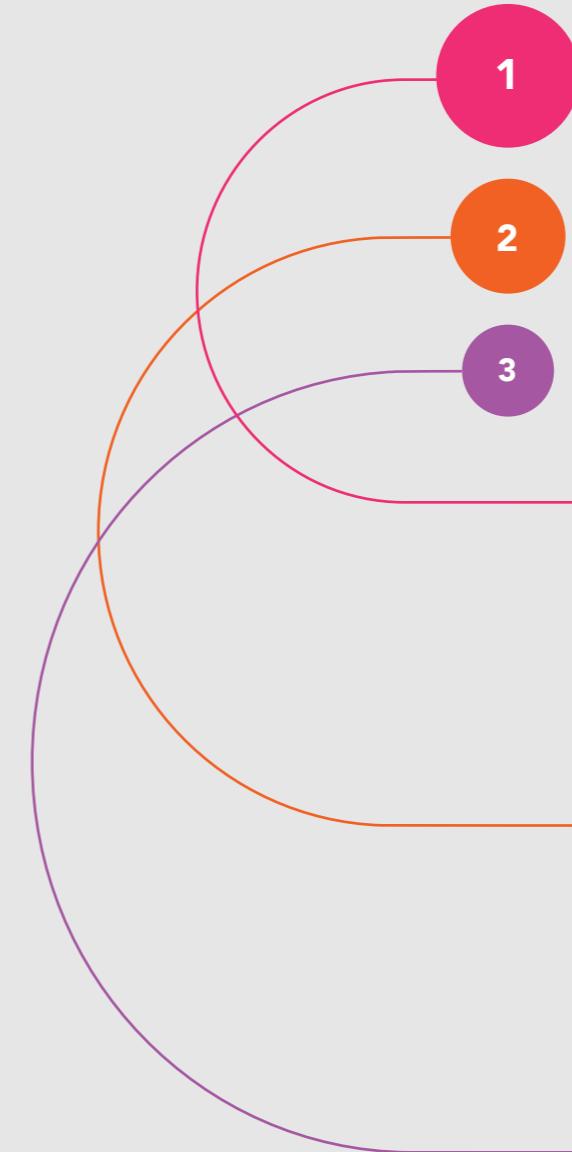
## Tactics and Execution

If there's a single truism that spans just about every major industry, it's that the industrial reality is far more complex than it seems on the surface. This is especially true when it comes to the future of data and its quickly evolving role in digital maturity.

As we conclude this chapter, hopefully we've equipped you with a new means to evaluate and measure progress throughout your journey. A serious consideration of digital maturity is an essential step on the road to implementing and realizing the benefits of Industrial DataOps.

**The following chapters will move us into that field, starting with a look at DataOps as a discipline.**

## Last Word: Digital Maturity



### Stay focused on the long-game:

Remember that the ROI curve is exponential, not linear. There will be gains along the way, but the key is to build enough momentum through tools and processes so that digital becomes effortless across a broad range of stakeholders.

### Measure holistically in addition to a per-project basis:

Individual projects absolutely need to be justified and reported on, but success (or failure) in one project isn't always representative of gains in digital maturity.

### Digital maturity spans people, process, and data:

It is very much a multidimensional means of change management. We recommend evaluating on Purpose & Strategy, People & Mindset, Value Capture Process, Information Architecture, and Data Ubiquity.

---

## Chapter 2

# DataOps as a Discipline: Extracting Value From Data



# From Digital Maturity to Industrial Transformation

For all industrial organizations, the intelligent use of data produced by operational technology (OT) systems is central to industrial organizations' efforts to improve operational excellence. So while everyone is talking about their digital transformation, use of data, scaling and time to value, very few in the industrial world are actually reaping the benefits.

And it isn't data that is the challenge. OT data is the raw material that enables organizations to build more efficient, more resilient operations and improve employee productivity and customer satisfaction. This OT data is available in abundance, but industrial organizations struggle to generate value from their increasingly connected operations—with IDC showing that only one in four organizations analyzes and extracts value from data to a significant extent.<sup>9</sup>

## The Challenge:

Only one in four organizations extracts value from data to a significant extent. Data dispersion and a lack of tools and processes to connect, contextualize, and govern data stand in the way of digital transformation.

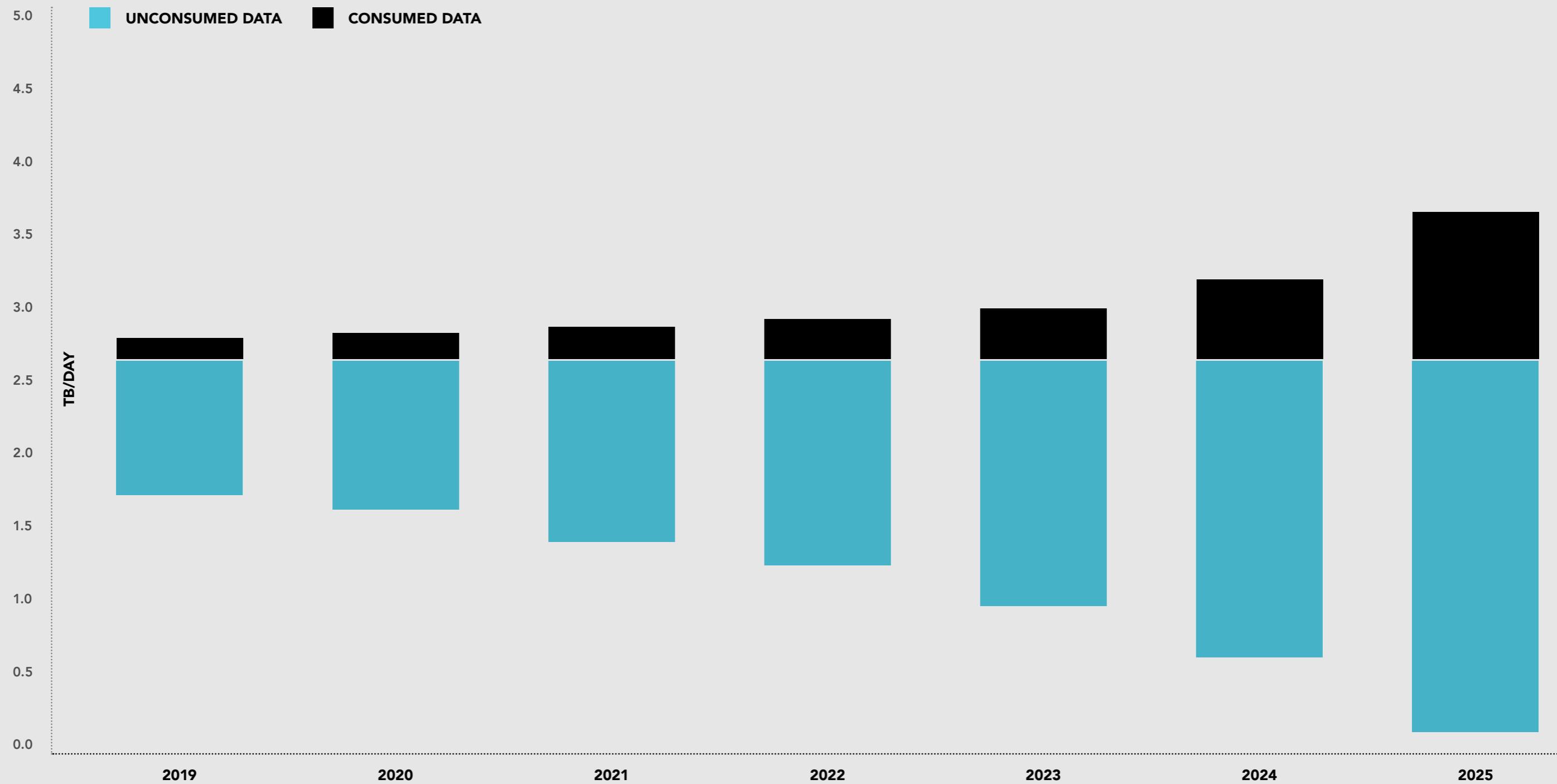
## The Opportunity:

Industrial DataOps promises to improve time to value, quality, predictability, and scale of the operational data analytics life cycle. It's also a stepping stone to a new way of managing data within the wider organization, enabling it to cope with growing data diversity and serve a growing population of data users.

The lack of appropriate tools and processes are significant obstacles resulting in data workers spending almost 90% of their time searching, preparing, and governing data.<sup>10</sup> A fear of missing data value often led organizations to prioritize data centralization over data organization. In turn, this led to poorly thought-out "data swamps" that only perpetuate the issue of dark and uncontextualized data.



**Figure 5: Data Generation and Consumption in a \$250 Million Industrial Operation, 2019–2025**



# DataOps as a Discipline

Before considering Industrial DataOps in detail, it's worth taking a look at how the early adopters of DataOps in industries such as banking, retail, and pharmaceuticals, responded to the challenges of operationalizing data.

While heavy-asset industry faces many specific challenges, the broad challenge around operationalizing data for value is shared across many sectors. This universal need is what gave rise to DataOps as a discipline, and saw it gain traction in a range of domains.

The IoT boom brought data, and the promise of data, to the forefront of business strategies across the globe. As the world ran towards the potential of data to drive meaningful change, DataOps emerged as the leader for operationalizing data in enterprise.

To refer back to the Forrester definition from Chapter 1, the power of DataOps lies in its ability to:

**"Enable solutions, develop data products, and activate data for business value across all technology tiers, from infrastructure to experience."<sup>12</sup>**

DataOps is a methodology for industrializing data management and the data analytics value chain. It applies automation, agile methods, to the data life cycle, improving time to value, quality, predictability, and scale of data analytics.

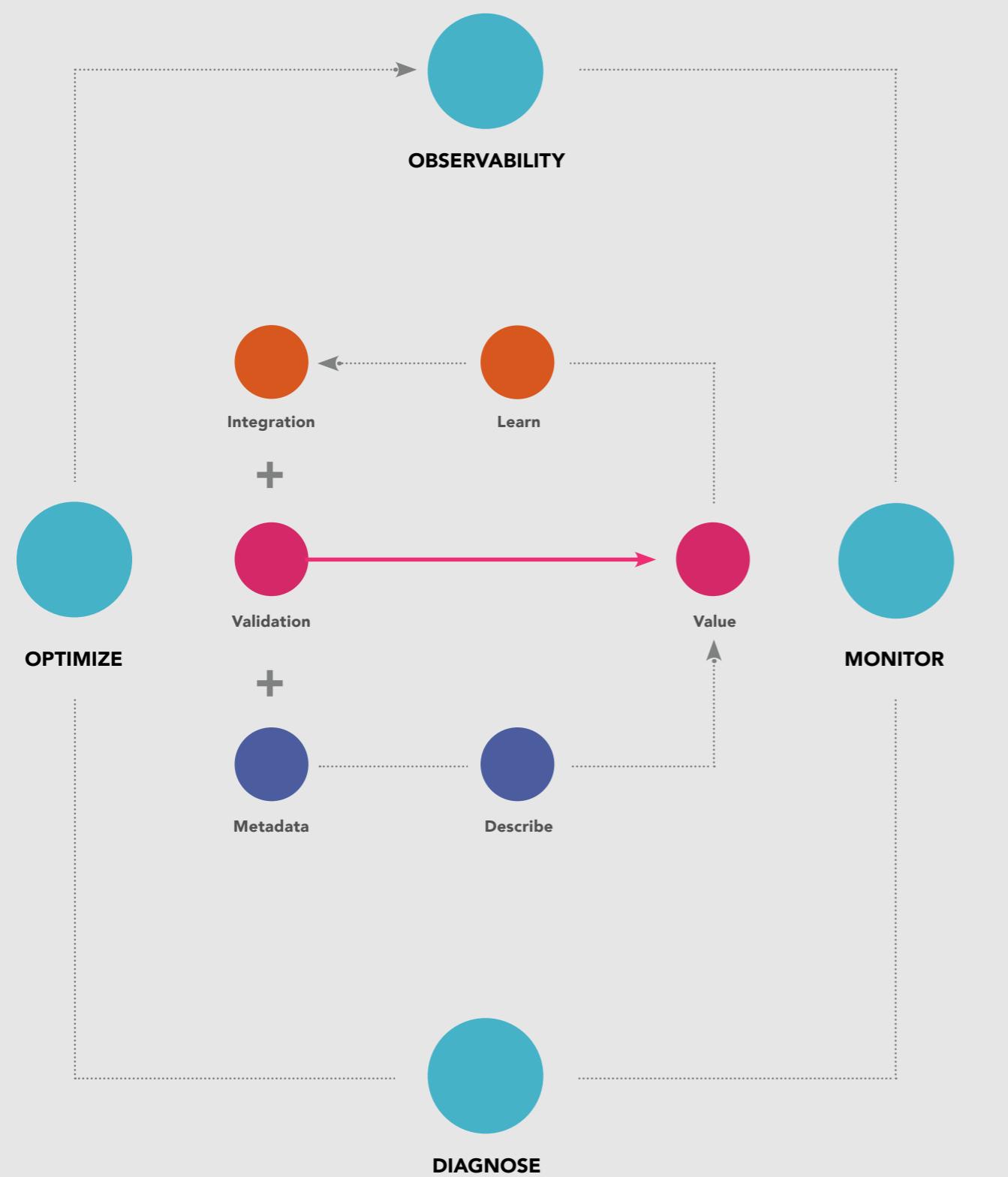
**"By 2023, 60% of organizations will have begun implementing DataOps programs to reduce the number of data and analytics errors by 80%, increasing trust in analytic outcomes and efficiency of Gen-D workers."<sup>13</sup>**

DataOps platforms help data workers deploy automated workflows to extract, ingest, and integrate data from industrial data sources, including legacy operations equipment and technology.

DataOps offers a workbench for data quality, transformation, and enrichment, as well as intelligent tools to apply industry knowledge, hierarchies, and interdependencies to contextualize and model data. This is then made available through specific application services for humans, machines, and systems to leverage.



**Figure 6: How Components of DataOps Interact to Continuously Deliver Business Value**



# Direct and Indirect Benefits of DataOps

## Efficient Data Management

DataOps maximizes the productive time of data workers with automated data provisioning, management tools, and analytic workspaces to work with and use data safely and independently within specified governance boundaries. The approach can be augmented with AI-based automation for various aspects of data management—including metadata management, unstructured data management, and data integration—enabling data workers to spend more time on use case development.



## Improved Data Accessibility

For many organizations, the current use of data is limited by dispersion across silos, spotty integration, and accessibility for centralized applications. Even where data sources are connected, data often lacks context due to limited documentation at the data's origin or information loss due to inconsistent structure or tagging.

DataOps technology uses AI to enable rapid ingestion and contextualization of large amounts of data. And by improving data accessibility, DataOps brings a paradigm shift in how the organization accesses business-critical information, improving decision-making quality, reducing risk, and lowering the barriers to (and skills for) data innovation.

## Rapid Development of Use Cases and Application Enablement

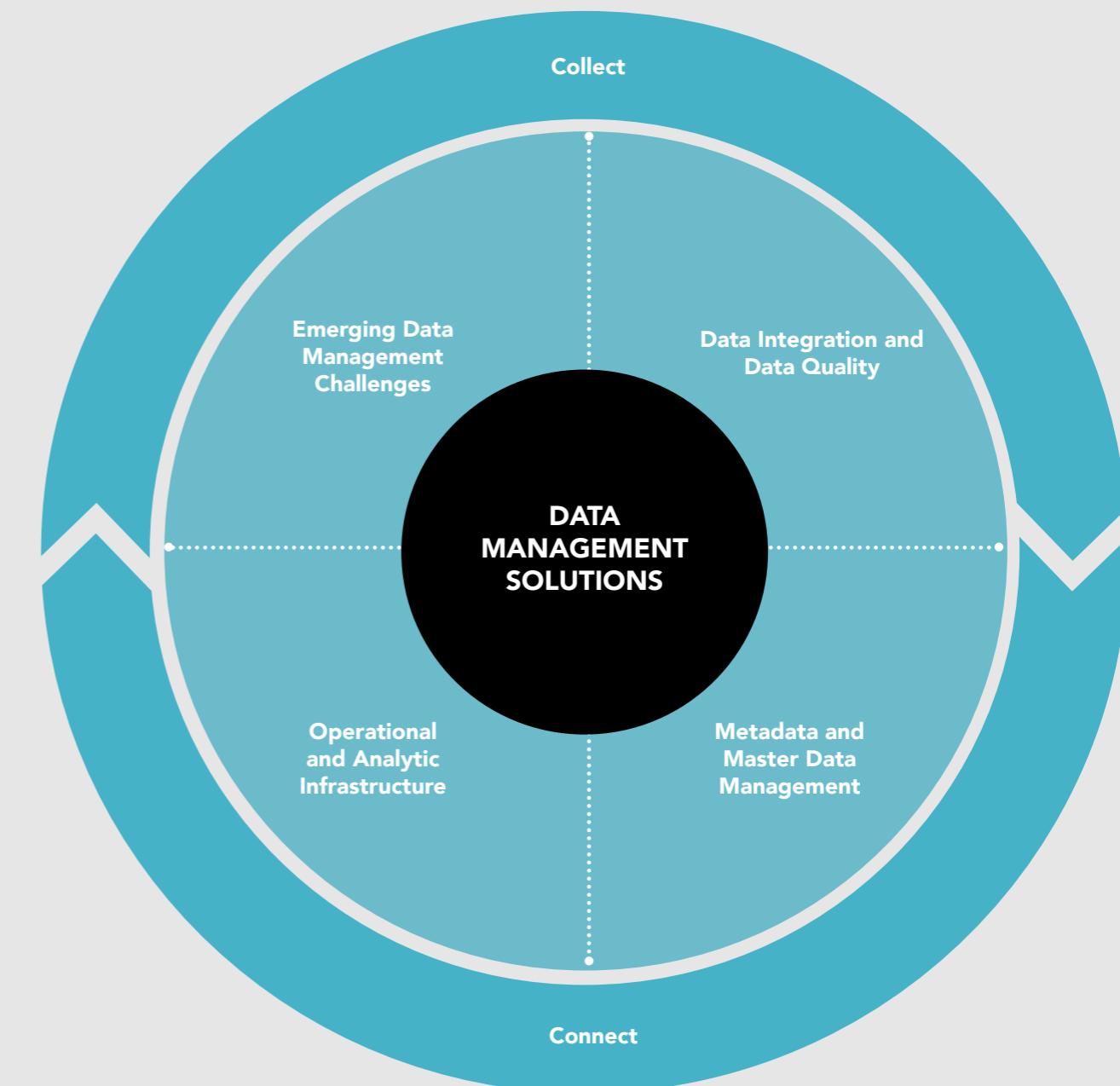
DataOps aims to shorten the time to value of data by making proofs of concept (POCs) quicker and cheaper to design, offering tools to operationalize and scale them.

## Enterprise Data Governance as a By-Product

DataOps enables companies to set and enforce the basic principles for managing data. If implemented successfully, the approach provides consistency and ROI in technology, processes, and organizational structures, with better operations data quality, integration and accessibility, and stewardship. A DataOps platform should also enhance data security, privacy, and compliance with tracking, auditing, masking, and sanitation tools.



**Figure 7: Data Management Solutions**



# Industrial DataOps Can Deliver Untapped Value for Heavy-Asset Enterprises

---

**DataOps is the clear frontrunner to become the driving force for transformation in industry.**

The definitions emerging in DataOps as an overall discipline have meaning across the business landscape but have a particular importance in industry. Not only do the traditional data challenges felt across all organizations become much more weighty in industrial sectors (both in terms of operational consequence and inherent data complexity), but industry also has its own set of unique challenges.

The scale and scope of these differences are such that a set of specific principles, tools, and techniques are necessary to successfully operationalize data in industry. This novel approach is what has been termed **Industrial DataOps**.

**Data must be made available, useful, and valuable in the industrial context.**

Amid the rush for change in the industrial world, the promises of data quickly became disappointments. In a push to show digital execution, many have embraced the AI hype.<sup>14</sup> This has led to quickly demonstrable digital POCs, yet is failing to yield truly operationalized—and even less scaled—concrete business OPEX value.

To adequately extract the value of industrial data insights, it's essential to make operationalizing data core to your business strategy.

This translates into developing and scaling mission-critical use cases across safety, efficiency, and sustainability. Data must be made available, useful, and valuable in the industrial context. We'll now consider the key steps, opportunities, and challenges associated with deployment of DataOps in an industrial organization. This is the route to extracting full value from your data.

## Step 1:

# Making Industrial Data Available



The fastest path to tapping into the value potential of digitalization in industry starts with getting the right data to the right user, with the right context, for the right problem, at the right time. However, the reality is that industrial data is still not easily accessible. It remains trapped in different systems, requiring data scientists to spend hours searching for it, collecting it together meaningfully, and preparing it for analysis.

## Trust in Data

Industrial enterprises need to be able to trust the data that they are putting into operations, simply because the cost of failure is too great.

**"We can't operationalize unless we trust the data. If something fails and we can't provide auditability we are finished as an industry."**

Stig Sund, Digital Operations Manager, Aker BP

Industry faces a particularly difficult challenge considering the nature of the data being analyzed. This data goes beyond standard tabular enterprise data. The bulk of data collected from operations revolves around process and instrument data (P&ID) and requires subject matter expertise to make sense of it.



In addition to being highly specific to industry, the time-series data collected is fraught with challenges of its own. Anyone dealing with streaming data understands that data quality and consistency are major thorns in the side of data scientists. Data quality issues become hard to check and must be done continuously, highlighting the importance of constant monitoring.

## IT/OT/ET/X Data Convergence

Traditional data silos are an obstacle to extracting value from data. Companies can make the most of Industrial DataOps to begin the integration process that spans asset and data life cycles across IT, OT, ET (engineering technology) and X (other types) data. The resulting converged data will support resilient decision-making across the organization and unlock the potential of fully fledged digital twin applications.

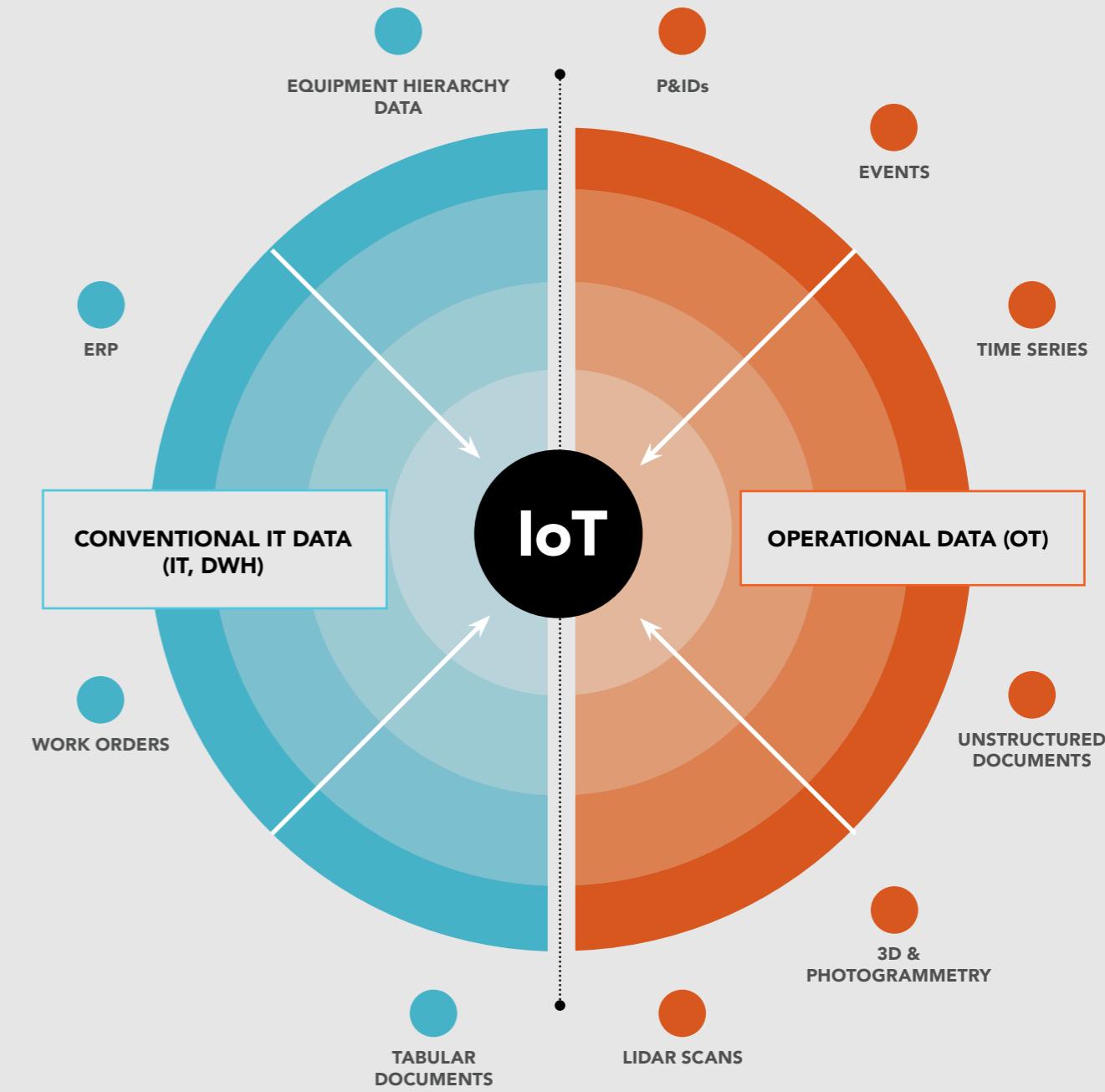
## Advances Across Technologies Make IT/OT/ET/X Data Convergence More Possible Than Ever



IT/OT/ET/X data convergence is not about turning IT pros into plant engineers or machine operators into data scientists—although the latter is indeed happening regardless. It is instead about executing on a strategy to align and bring together formerly isolated subject matter experts (SMEs), cultures, platforms, and data deployed by OT and IT teams to improve operational performance through unified goals and KPIs.

In practice, this is happening through the adoption of a new data and digital platform that contextually fuses IT, OT, ET and other data types such as audiovisual part of X data making this contextualized data conveniently available to a growing audience of data consumers, both inside the enterprise and across its partner ecosystem.

**Figure 8: Relationships Between IoT, Conventional IT Data, OT Data, ET Data, and Other Data Types (X Data)**



## Step 2:

# Making Data Useful

Industrial data becomes truly useful (read: fit for operational purpose) when it is integrated, contextualized, and made securely available, explorable, and actionable to all data consumers (human and machine) within and outside the industrial enterprise. This should encompass all the various sources and formats including sensor data, process diagrams, 3D models, event histories, asset models, and unstructured documents.

## Contextualization

Data contextualization involves connecting all the data for a clearer understanding of an asset or facility. This should be at the core of an Industrial DataOps platform, establishing meaningful relationships between data sources and types to help users find and utilize relevant data from assets across the operation.

A petroleum engineer, for example, would understand the sensor data streaming from an electric submersible pump of an oil-well site, but a data scientist might not. Contextualization links the pump identity from the asset hierarchy to its sensor data and related work orders and connects it to the asset's 3D model.

Similarly, in the steel industry, a data scientist might not be able to grasp the complexity of predictive quality and steel-grade monitoring without a solid knowledge of the underlying chemicals and laws of physics. If they are given more context with a 3D model or knowledge graph, however, they would be able to visualize the operational context to develop models and data applications, in this case for anomaly detection.

## Further Adding to the Complexity

There are challenges inherent to operating in large, siloed organizations of industry: among them the different reporting lines, varied analytics workflows, competing business interests, and varying incentives. If industry is to break free of the complexity, tools and technologies must be the driver of change.

Adding insult to injury, the rise of AI/ML, combined with the difficulty of finding data scientists, is imposing its own set of requirements on data modeling, data source availability, data integrity, and out-of-the-box contextual metadata. These requirements are often very differentiated from those of the conventional BI user.

Data engineers working on industrial digitalization projects struggle with access to key source system data, in a way that is reminiscent of the year 2010 in non-industrial verticals. Industrial companies are not only facing the same challenges as, for example, their retail peers, but are presented with a superset of challenges resulting from the IT/OT convergence, and associated non-conventional IT-only data velocity, variety, and volume.

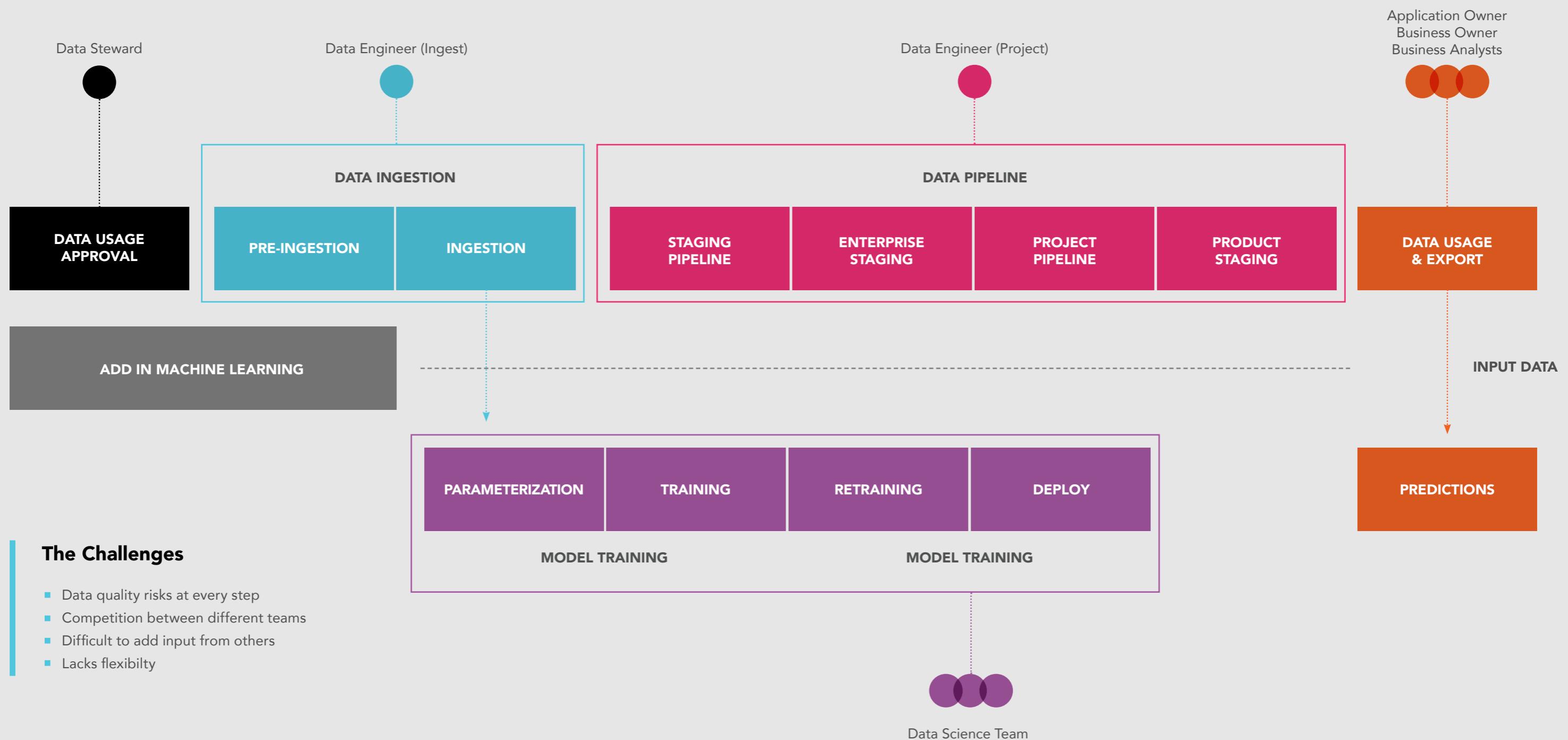
The challenges of making data useful, compounded by the complexity of industrial data, require special solutions. Contextualization is core to being able to meet these business challenges and deliver scalable operational outcomes in industry. Indeed, contextualization forms an important foundation to be able to automate the journey of data to solutions. Without contextualization, data lacks meaning and is poorly suited for Industrial DataOps.

## Enabling Industrial Hybrid MLOps

Industrial DataOps platforms offer the combination of data-driven statistical and physics-driven process modeling and simulation. While each approach has its pros and cons, often an ML model based on a hybrid of the two will provide the best results. These tools empower developers with workflows compatible with third-party AI tools and other necessary tools to develop, train, and manage hybrid ML models. This enables them to operationalize use-case-specific data subsets efficiently and at the desired scale.



**Figure 9: A Typical Data Pipeline for Analytics With Associated Workflow Challenges**



## Step 3:

# Making Data Valuable

Extracting maximum value from data relies on being able to apply advanced models to produce insights that inform optimal decision-making, empowering operators to take action with confidence. This, in a nutshell, is what is meant by operationalizing data into production for value.

Advanced models combine data science with physics to generate synthetic data and advanced insights. This is complemented with machine learning and deep learning for scale. It's crucial to observe, analyze, and optimize to deliver reliable forecasts and actionable insights.

## Democratizing DataOps

Industrial DataOps platforms enable data users with low-code or no-code application development and model life-cycle management tools. This democratizes DataOps and facilitates a more collaborative working model, where non-professional data users can perform data management tasks and develop advanced analytics independently within specified governance boundaries. This democratization of data helps store process knowledge and maintain technical continuity so that new engineers can quickly understand, manage, and enrich existing models.

## Going Beyond Proof-of-Concept

Too often, digital operation initiatives get trapped in "POC purgatory", where scaling pilots takes too long or is too expensive. What holds them back are the IT/OT and OT/data science divides, and the inability to produce and access contextualized, quality data at scale.

By connecting data users with disparate operational data sources, an Industrial DataOps platform helps bridge those divides on the path to use-case operationalization. ML libraries of standard industrial use cases help developers save time when collecting data and developing and training their models. Data scientists can leverage this library and use it with component-level data. Once a use case is developed and the outcomes are satisfactory for one component of the plant, the contextualization of asset data allows it to be scaled to plant or fleet level.

Some examples of common use cases in heavy-asset industries are: maintenance workflow optimization, engineering scenario analysis, digitization of asset process and instrumentation diagrams (P&IDs) to make them interactive and shareable, and 3D digital twin models to support asset management.



# What to Consider When Adopting Industrial DataOps

---

**Asset-intensive organizations should look to Industrial DataOps to unleash the full potential of IT/OT/ET/X data and to transform their traditional operating model. When starting on this journey, companies should:**

Think of AI as a critical tool for both fact-driven decision-making and efficient management of the data supporting it. Bypassing human “midstream” data handling is key.

“Data liberation” is critical to getting full value from DataOps. Maximizing your data extraction capabilities will make it easier to realize DataOps with your existing IT and OT architecture, limiting the need to invest in additional systems integration and OT data sources.

Develop a strong data governance model for IT/OT/ET/X data. This will dictate how new data is connected and integrated into the overall data architecture. It will also help serve a growing population of data and analytics business users.

Prioritize data organization over centralization. Start driving the connection and mapping of all relevant data sources with a clear list of target use cases in mind.

As part of the governance model, all new data sources must have a connection, tagging, sharing, and integration plan. Note that not all DataOps platforms have the same capabilities. Alignment with your goals, industry track record, and domain expertise should drive selection criteria. (See Appendix for a complete guide to provider evaluation).



# Moving From Theory to Practice

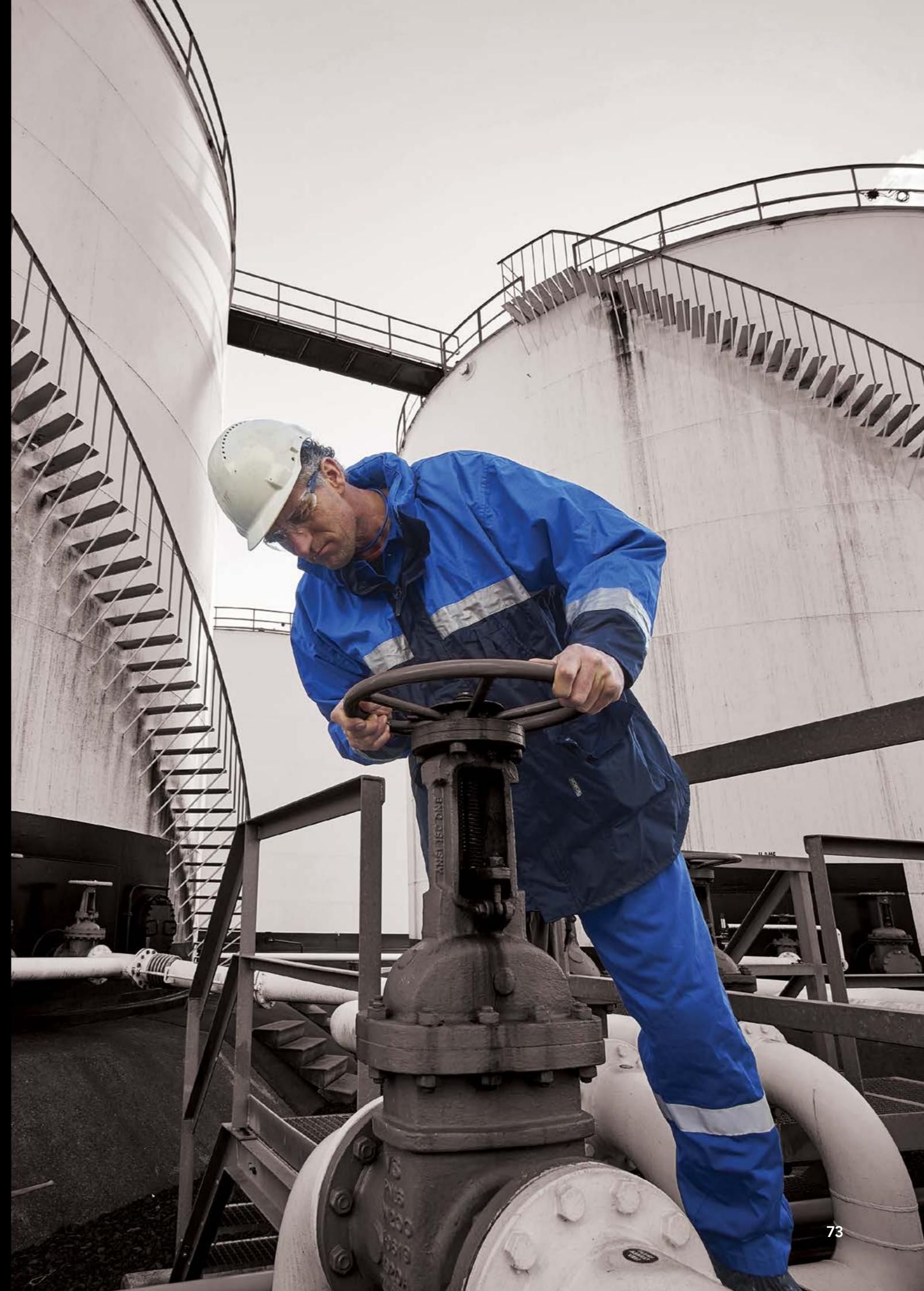
As discussed in previous chapters, Industrial DataOps requires your organization to take some critical steps to start the journey on the right foot. For starters, it requires stronger cohesion among data stakeholders. Data science and IT must collaborate well beyond data access and resource allocation, while business should be involved in data projects well beyond the typical demand and validation stages.

Organizational divides compound a company's inability to access data at scale, and make asset analytics pilots too long or too expensive to operationalize. Bridging these organizational and operational gaps is a balancing act that requires focus and leadership.

**Deploying the right tools—feature-rich, intuitive, and easily scalable—can be a catalyst for lasting, positive change.**

# Principles of Industrial DataOps

## Chapter 3





**“DataOps is the ability to enable solutions, develop data products, and activate data for business value across all technology tiers from infrastructure to experience.”**

Forrester<sup>15</sup>

Now we come to the core principles of the novel Industrial DataOps approach. This chapter will lay out a set of essential concepts to guide you on your way to extracting maximum value from your data.

Some readers may be familiar with long lists of general DataOps principles, such as those provided in the formidable DataOps Cookbook.<sup>16</sup> You may have also seen some of the detailed DataOps evaluation criteria offered by forward-thinking technology consultancies (Fig 10).

**Figure 10: General DataOps Principles From the DataOps Cookbook**



15. Goetz, M et al (2020). DataOps For The Intelligent Edge Of Business: Go Beyond Traditional Data Management And Governance. [Online]. Forrester. [11 Mar 2021]. Available from: <https://www.forrester.com/report/DataOps+For+The+Intelligent+Edge+Of+Business/-/E-RES162717>

16. Bergh, C. The DataOps Cookbook. [Online]. (2nd ed.). Published online: DataKitchen. [23 Mar 2021]. Available from: <https://datakitchen.io/the-dataops-cookbook/>

**Figure 11: Example of Generic DataOps Evaluation Criteria**



While much of the general DataOps material provides useful background information, the purpose of this book is to provide specific, practical guidance to help organizations in heavy-asset enterprises operationalize data for value.

Since our focus is on Industrial DataOps, the principles we put forward here are squarely from the perspective of digitalization leaders focusing on industrial operations.

# 7 Principles of Industrial DataOps

## 1. Don't do POCs

POCs are worthless to your CFO. Every digitalization use case you start needs to be architected to run live in production, at scale, before the first line of code is written.

## 2. Think in data products, execute in data domains

For data to be operationally useful at scale and for critical operations, it needs to be productized. To productize your data, focus on the most valuable operational data domains first, not on the enterprise-wide master data landscape all at once.

## 3. Your data needs to speak human

The key to creating value from data lies in data context and interpretability by data consumers in business operations, not on the collection of more data.

## 4. Business technologists are your digital heroes

Your target persona for data product manager as well as data consumer is your business technologist, not your data engineer.

## 5. Autonomous industry is your north star

The goal of Industrial DataOps is autonomous industry, not universal data availability or better data engineering.

## 6. Old technology stacks do not work

Executive support, a transformation mindset, and upskilling of people are needed. Without the right new tooling however, they all remain impotent.

## 7. Quality is king

Speed matters, but running fragile solutions leading to failures in production simply does not work in industry.

1:

# Don't Do POCs

**Proofs-of-concept (POCs) are worthless to your CFO. Every digitalization use case you start needs to be architected to run live in production, at scale, before the first line of code is written.**

**Figure 12: Value Capture Reality**

Despite supportive technology trends in abundance, data activation remains stuck in POC purgatory.



Most of us have been there, or are there right now: in POC purgatory. It's easy to understand why. Over recent years, we have seen a well intended, though at times obsessive, bottom-up and top-down drive to show digitalization use cases to all industrial enterprise stakeholders, not least to public markets, industry peers, and trade press.

By implementing digitalization use cases following a systematic platform approach rooted in best Industrial DataOps practices, enterprises can break free of the POC purgatory cycle and focus on use case innovation that delivers in production at scale.

## The optics trap of quick digitalization success has trumped genuine

This has unfortunately often resulted in prioritization of easy-to-communicate digitalization showcases rather than actually delivering in-production operational digitalization solutions of real bottom-line value. The optics trap of quick digitalization success has trumped genuine in-production success.

It is important to note in this context that POCs themselves are not to blame. It makes sense to validate use cases before their operational scaling. Challenges, and consequent loss of real value capture, emerge when operational scaling is not intrinsic to the overall digitalization framework, technology architecture, and processes followed. In other words, before a POC is signed off, the execution architecture needs to account for its operational production environment scaling, and not just its theoretical ROI viability.



2:

# Think in Data Products, Execute in Data Domains

**For data to be operationally useful at scale and for critical operations, it needs to be productized. To productize your data, focus on the most valuable operational data domains first, not on the enterprise-wide master data landscape all at once.**

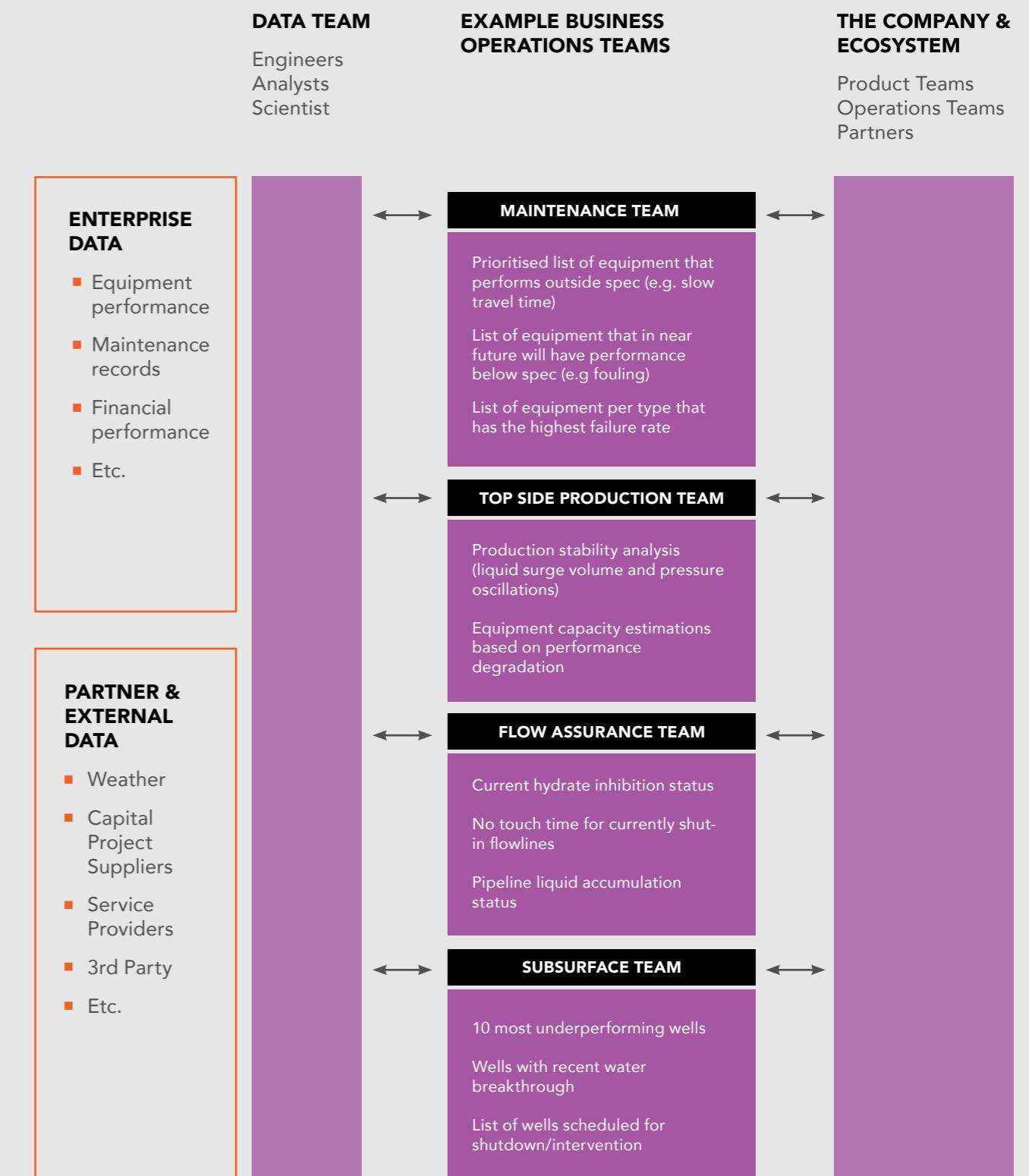
Advances in cloud data storage and elastic processing have catapulted us into the early era of always-on secure data access for all data consumers within and around the enterprise.

That said, there remains an exponentially growing need for a similar step change to happen at the intersection of data engineering

and custodianship with the business expertise for that data domain.

This shift from data availability to data products as a service (Fig. 13) is what will allow us to transform our data swamps into operational data architectures of real business value.

**Figure 13: Data-Products-as-a-Service Model**

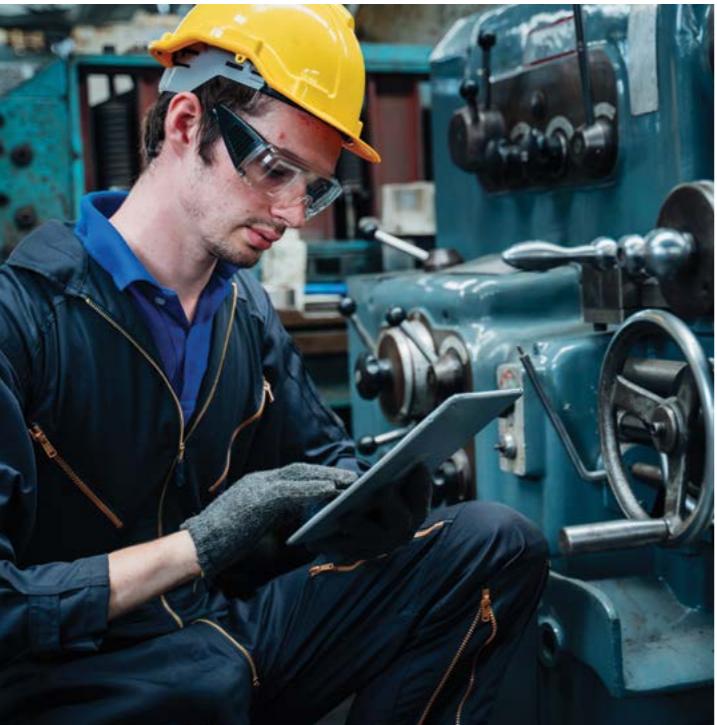


Any organization hoping to successfully transition into offering data products as a service, needs to embrace a shift in data product ownership. The necessary move is from a centralized data team, such as digital or data center of excellence, into a collaborative setup, where each data domain is co-owned by the

respective business function producing the data in their primary business tools. After all, it is the business operations team that understands the data in context best, and is therefore best placed to communicate and provide data products as a service to other data consumers.

## What is a Data Product and What Defines a Data-Products-as-a-Service model?

1. Data products are a team sport: The data team partners with business operations to tackle specific problems using data.
2. Data products have an owner, support, SLA, and clear definition.
3. Data products have an SLA from the entire data domain team, not just the data engineer.
4. Data-products-as-a-service flow is bi-directional, from the domain data team to the company and back.
5. Domain expertise is blended directly into the data products themselves.
6. Data product team members have more business functional experience for their data products, and are responsible for providing insight as opposed to rows and columns.
7. Data-products-as-a-service fuses a service-oriented business partnership focusing on value with a product-oriented SLA focusing on reliability and data quality.



To successfully implement Industrial DataOps, it's essential to move from a conventional centralized data architecture into a domain data architecture (or data mesh).

This solves many of the challenges associated with centralized, monolithic data lakes and data warehouses. The goal becomes domain-based data as a service, not providing rows and columns of data.

**For domain data architecture to work, the data product owner teams need to ensure their data is discoverable, trustworthy, self-describing, interoperable, secure, and governed by global access control. In other words, they need to manage their data products as a service, not as data.**



3:

# Your Data Needs to Speak Human

---

**The key to creating value from data lies in data context and interpretability by data consumers in business operations, not in the collection of more data.**

Confronted with an exponential rise in data volume, velocity, variety, and value creation expectations, enterprises large and small are rushing to upskill their workforces to become better data consumers—or to be more data literate. Gartner formally defines data literacy as “the ability to read, write and communicate data in context,”<sup>17</sup> more informally expressed as “Do you speak data?” Data literacy includes an understanding of data sources and constructs, analytical methods and techniques applied to data, and the ability to describe the use case applications and resulting value.

In industry, most of our production and maintenance data consumers—those at the industrial AI revolution front line—are subject matter experts with deep domain, and domain data, expertise. But they are not data engineers, database specialists or solution architects. Nor should they have to be.

To deliver data literacy at scale and across data domains, it is our data management infrastructure that needs to rise to the occasion and accommodate the growing demands of the new data consumers. This acknowledgement lies at the heart of Industrial DataOps.

Instead of:

**“Do you speak data?”**

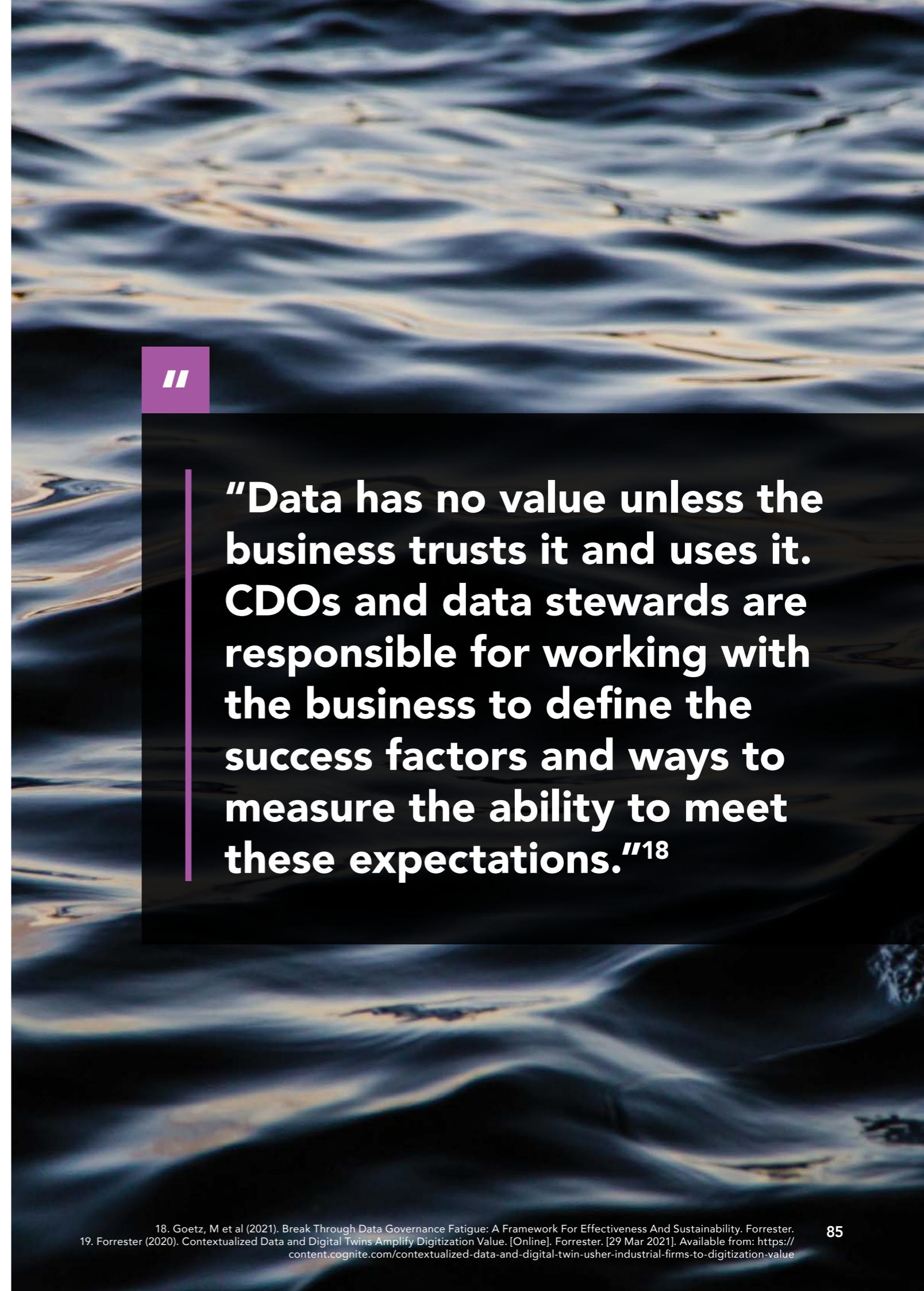
Consider:

**“Does your data speak human?”**

There is a growing realisation of this fundamental truth. Research shows that 79% of industrial leaders across IT and operations agree that data needs to become self-explanatory to data consumers without needing subject matter expert support.<sup>19</sup> In other words, it is the data that needs to speak human, not vice versa.

84 17. Gartner. Gartner Glossary. [Online]. [29 Mar 2021]. Available from: <https://www.gartner.com/en/information-technology/glossary/data-literacy>

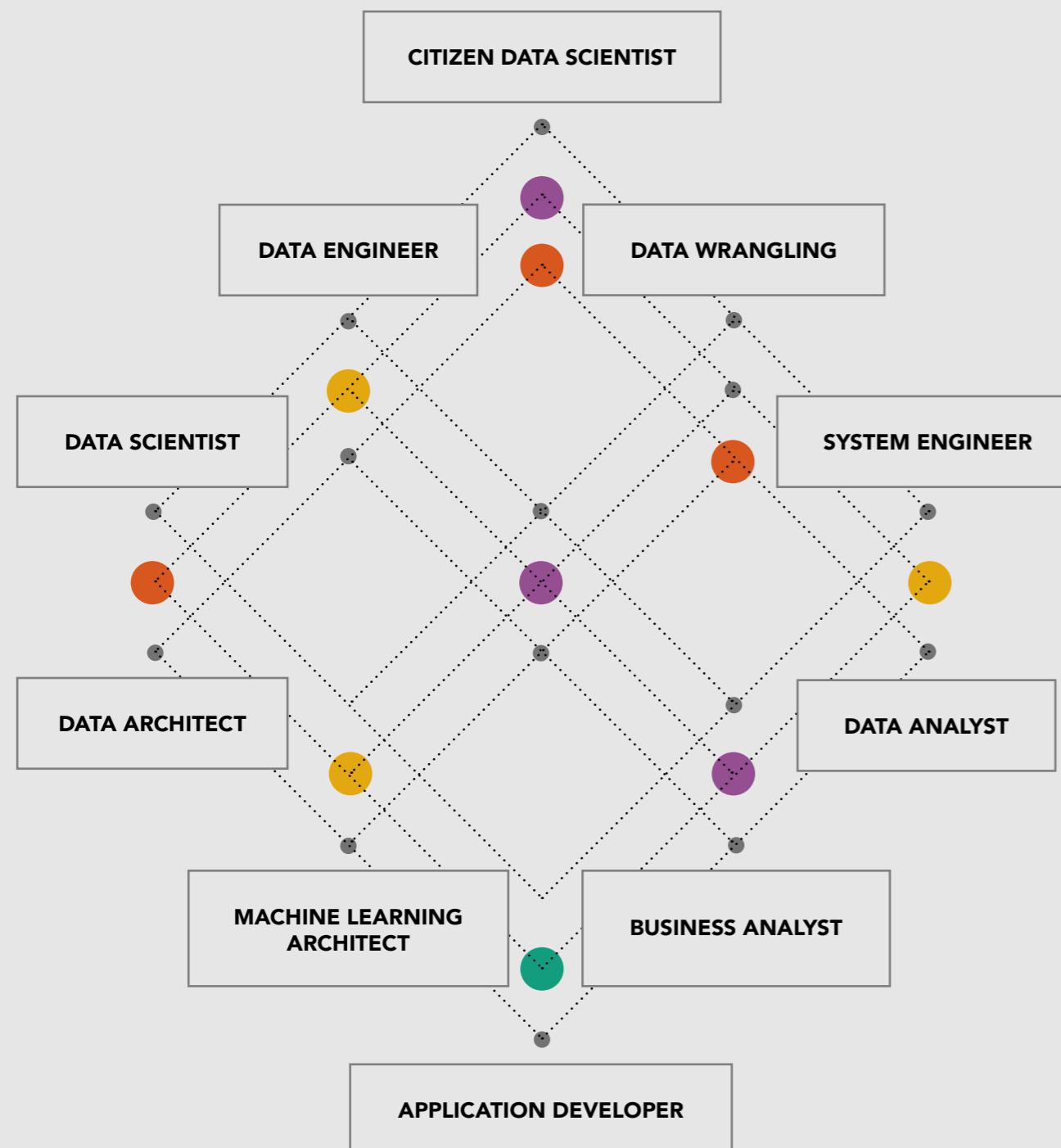
18. Goetz, M et al (2021). Break Through Data Governance Fatigue: A Framework For Effectiveness And Sustainability. Forrester.  
19. Forrester (2020). Contextualized Data and Digital Twins Amplify Digitization Value. [Online]. Forrester. [29 Mar 2021]. Available from: <https://content.cognite.com/contextualized-data-and-digital-twin-usher-industrial-firms-to-digitization-value>



“

**“Data has no value unless the business trusts it and uses it. CDOs and data stewards are responsible for working with the business to define the success factors and ways to measure the ability to meet these expectations.”<sup>18</sup>**

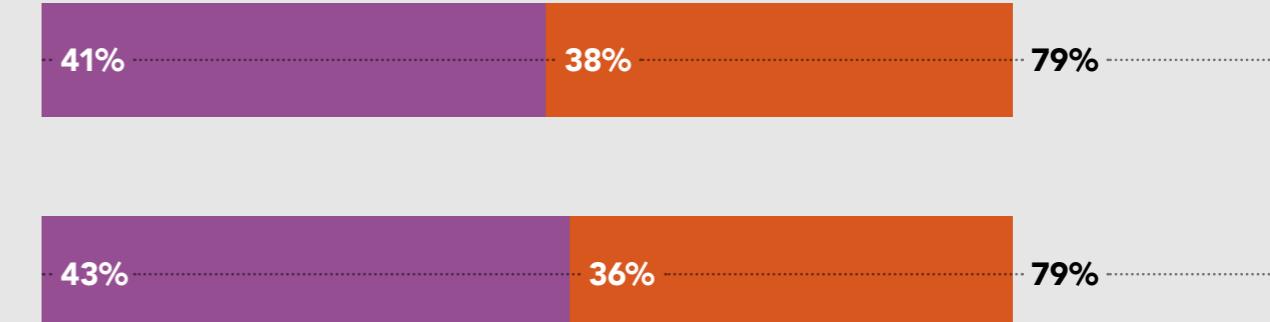
**Figure 14: The Importance of Data Accessibility Across the Organization**



Data needs to become self-explanatory to data consumers without needing subject matter expert support

A company is digitally mature when it can enable domain experts and other data consumers to do more with advanced data and analytics by themselves

● Agree      ● Completely agree



To achieve this level of true data literacy, it's critical to think in data products and to execute in data domains (see Principle 2. above).

It is equally important to fully embrace AI-enhanced active metadata curation, leveraging advances in neuro-linguistic programming (NLP), optical character recognition (OCR), computer vision, trained ontologies, and graph data models to facilitate increasingly automated IT/OT/ET/X data contextualization for intuitive human as well as programmatic data discovery and use.

**Industrials will see ten times higher returns from their investments in metadata than in data itself.**

Metadata can be defined as "any data that is used to enhance the usability, comprehension, utility, or functionality of any other data point."<sup>20</sup>

4:

# Business Technologists Are Your Data Heroes

Your target persona for data product manager as well as data consumer is your business technologist, not your data engineer. True digitalization happens outside your digital center of excellence. It happens in your core business operations by the actions of a new class of business-technology hybrids called business technologists.

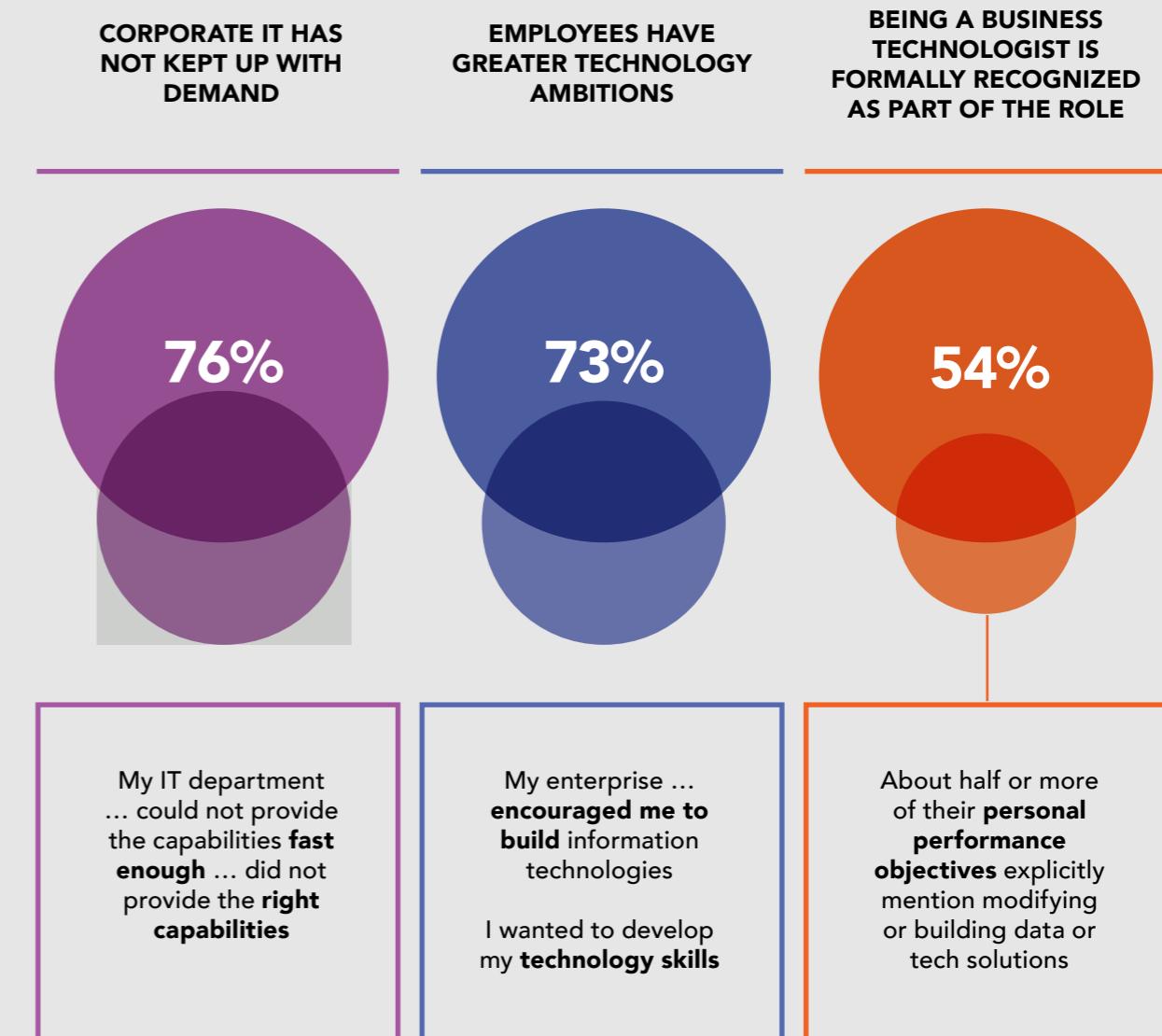
These line-of-business residents are already growing impatient with a lack of Industrial DataOps capabilities, which they need to support self-service discovery and data orchestration from multiple sources.

Such capabilities will enable them to make better decisions and increase their competitiveness. At the same time, industrial application developers crave trusted, consistent, and real-time data to build modern cloud-to-edge microservice applications using APIs, and they too increasingly reside within business operations.

**As new data consumer roles such as business technologists increase in popularity, corporate IT is under increasing pressure to deliver faster and more autonomous access to intuitive data products as a service.**

Self-service business domain empowerment is of course good news for the business of IT, and it offers an attractive recruitment route for domain data product collaborators as discussed in Principle 2. above (also see Fig. 11).

**Figure 15: Reasons Why Business Technologists Do What They Do**



Of course, the Industrial DataOps platform alone is not the full answer, not even with business technologists paving the way to the next segment of industrial data consumers. It needs to be complemented by enterprise data dashboarding and low-code application development capabilities. Fortunately, both are much more readily available across industries at present.

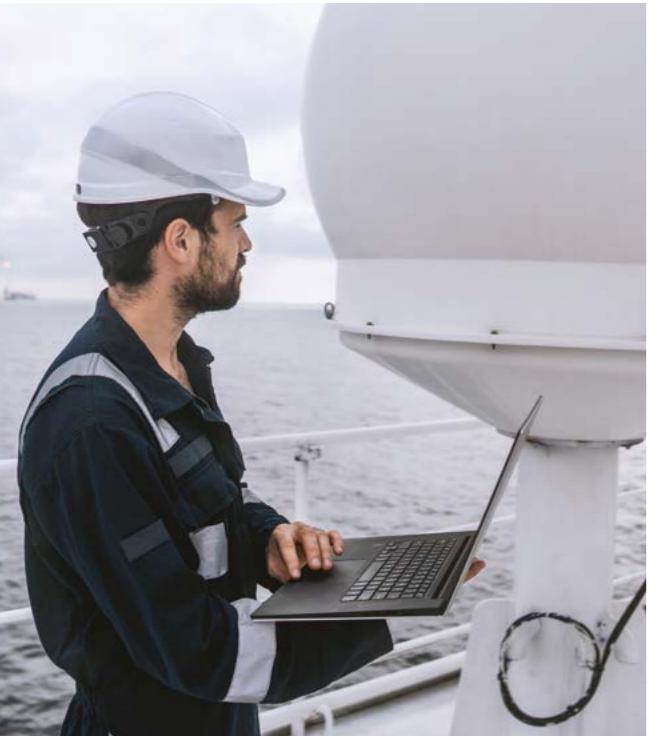
Last, to build trust and collaboration across IT and business, don't refer to your operations subject matter experts as citizen data scientists. They are domain experts. They can also be your best business partners for domain data product delivery as well as championing business technologists. What they are not is "citizens."

5:

# Autonomous Industry Is Your North Star

---

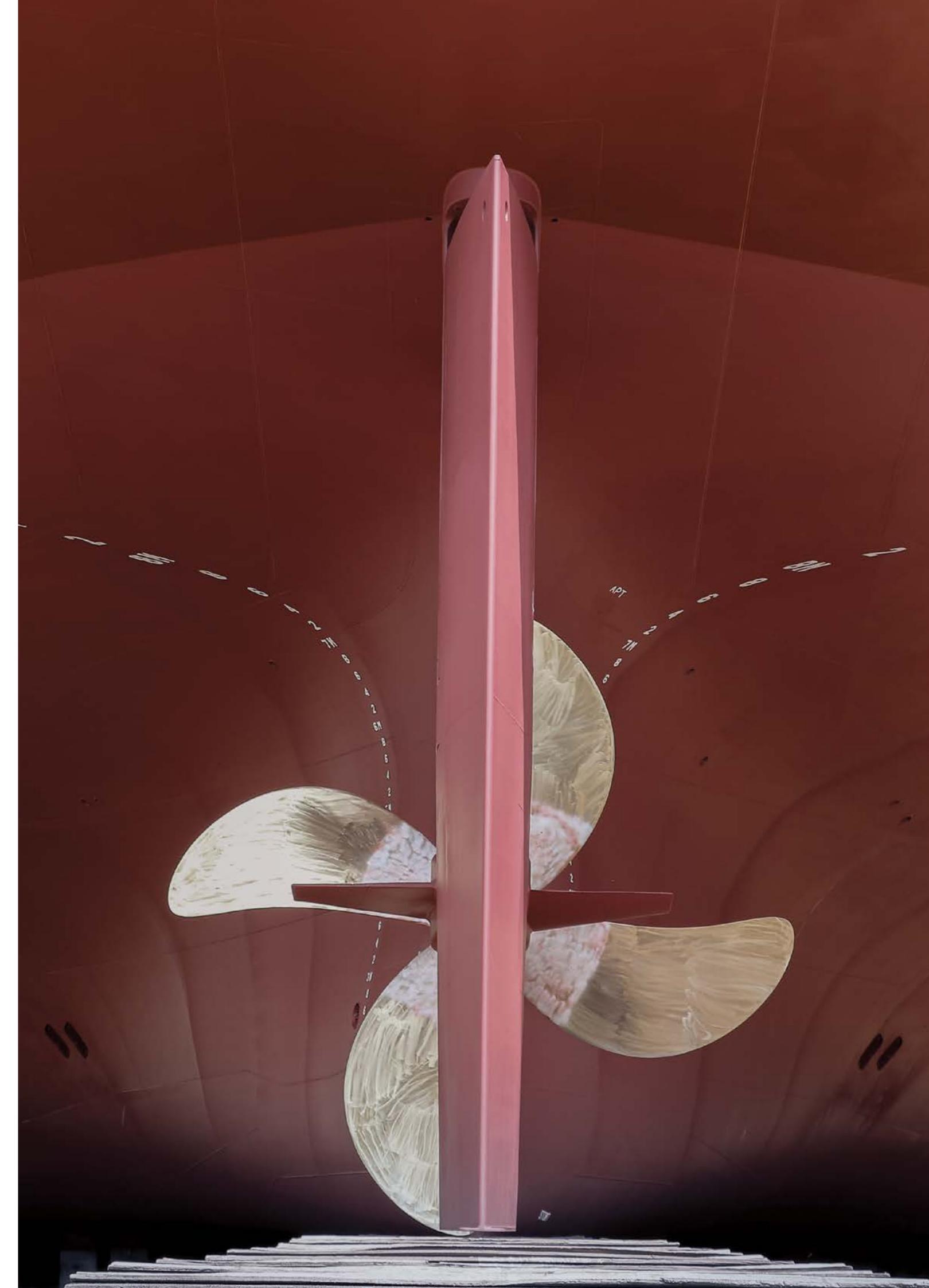
The goal of Industrial DataOps is autonomous industry, not universal data availability or better data engineering.



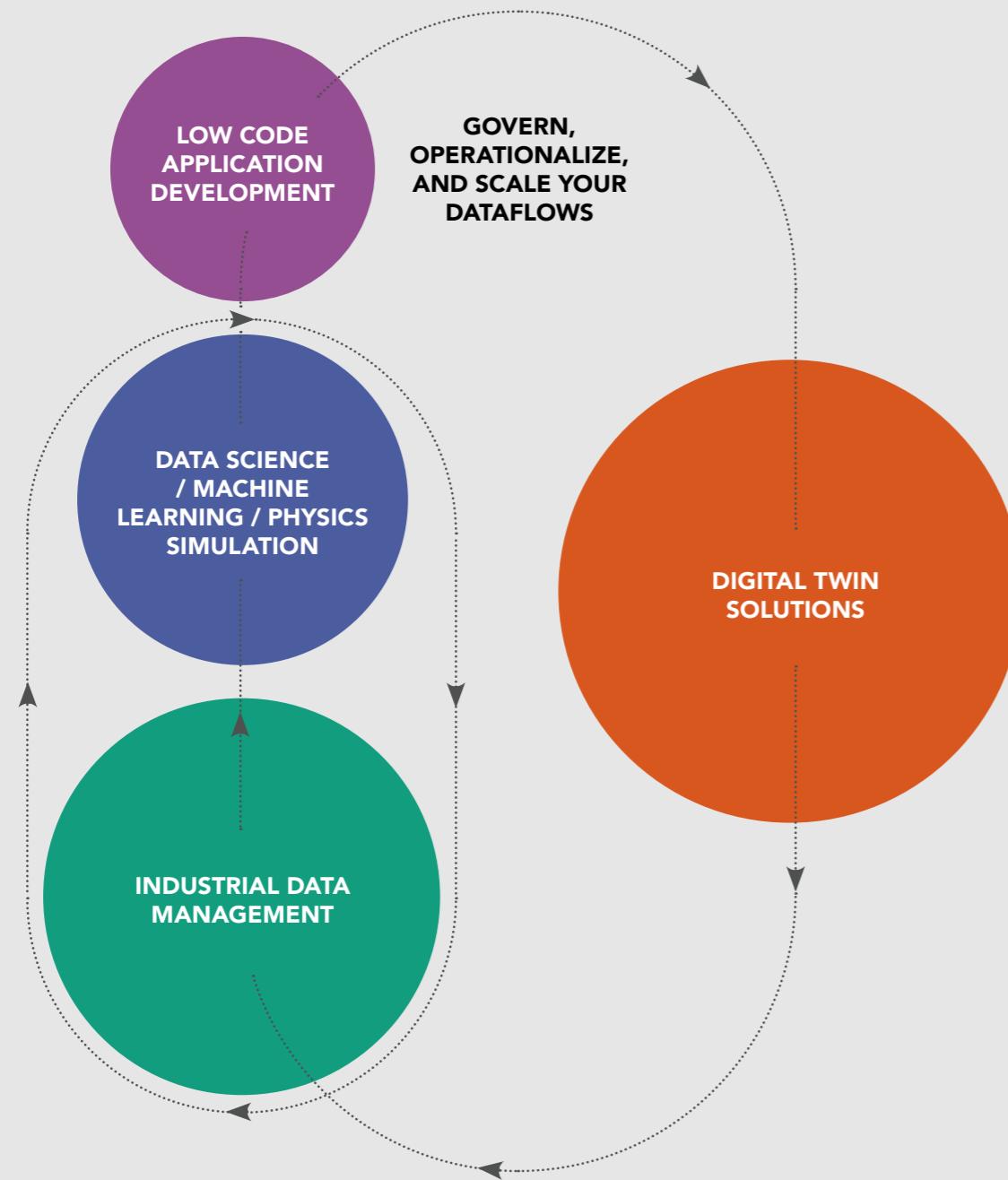
Industrial DataOps is not only timely and valuable now; it is foundational data infrastructure for the future.

This is because the data-driven operations capabilities it provides are equally necessary to achieve the goal of autonomous industry, the promise of Industry 4.0.

**This should always be your guiding target, your true north star.**



**Figure 16: Industrial DataOps Makes Industry 4.0 work**



**Without a robust foundation of Industrial DataOps, trust in data, data models, and data-driven recommendations remains low, resulting in inability to progress beyond the 1st or 2nd level of data-driven operations improvement.**

If fully autonomous decision-making with closed-loop intelligent production systems as the north star, paving our way there is already well underway, and requires largely the same core technology capabilities.



**DataOps is key to crossing the chasm from simply reporting data to data-driven operations:**

1. **Turn Data into Insights**
  - Visualize existing data and put it into context.
  - Users interpret information via graphs, dashboards, and make qualified decisions based on available data.
2. **Turn Insights Into Actionable Advice**
  - Enrich the existing data and create recommendation models with actionable advice.
  - From simple implementation of known equations to anomaly detection and machine learning.
  - Humans evaluate output and make qualified decisions based on the recommendations.
3. **Build Advice into Workflows**
  - Advice is delivered to the right persona via a defined process.
  - Users are prompted with advice about impending issues and how to proceed. Persona-based content can be delivered through virtual assistants, to provide status and actionable next steps.
4. **Automate Your Workflows**
  - Integrate models directly with selected systems to trigger automatic actions.
  - Enable closed-loop integrations with no human interaction.

6:

# Old Technology Stacks Do Not Work

---

**Executive support, a transformation mindset, and upskilling of people are needed. Without the right new tooling however, they all remain impotent.**

Although they affect the whole enterprise landscape, the disruptive potential of technologies such as AI, cloud computing, and inexpensive data collection is now disproportionately impacting many industrial sectors. These were traditionally insulated from disruption by heavy regulation, low threat of new competition, and extreme capital intensity. Not any more.

Power and utilities is perhaps the sector experiencing the greatest disruption. Organizations face a perfect storm of ESG and activist capital allocation, societal sustainability

pressure, renewable energies, distributed energy resources (DER), and the rise of electricity prosumers. Compounding these challenges, the sector is largely reliant on very outdated software technology architecture designed for closed site-level control in a largely static environment.

To not only adjust but to thrive, energy companies are aggressively rethinking their data and analytics architectures for the next decade, one set to be defined by constant change, collaboration, real-time data, and innovation.

Despite such profound business environment and technological change, there can sometimes be a tendency to downplay the role that new software technologies need to play, and resort to common executive phrases such as:

*"The technology is readily available, it is the change management and people organization that remains challenging."*

*"Before we start to address the technology side, we need to thoroughly assess and plan for cultural change implications."*

The idea that cultural change and mindset should be prioritized over technology is simply not true. As the change enablers disrupting our industries are all profoundly technology-led, we need to start with the technology—or address it in parallel at a minimum. Executive support, a transformation mindset, and upskilling of people are needed. Without the right new tooling however, they all remain impotent.

Changing processes or behaviors without the tools to do so—the concrete engines of change—is simply too abstract and disconnected, particularly from the engineering mindset. Technology disruption does, in fact, start with the technology.



So how do you go about finding the right technology stack to power your Industrial DataOps? Figure 15 provides a summary of key features that a domain data consumer-focused digital platform should provide.

In the Appendix (Industrial DataOps RFP Guideline) you will find a comprehensive list of questions to aid formal evaluation and selection of Industrial DataOps software.

The appendix covers the entire spectrum of considerations in detail, from use cases and past successes, to solution architecture, to security and software maintenance. It is intended as a complete guide and toolkit to support requests for proposals (RFPs) and similar procurement processes. Here, we have distilled this down to a core set of five key questions to get you started.

# Five Questions Industrial Companies Should Ask Before Purchasing Industrial DataOps Software



## 1. Have you aligned the internal needs of both your IT and operations organizations?

Industrial software finds its home at the crossroads of the IT and operations worlds. Any software purchase affects both. This means that it's essential to uncover the needs and requirements of both parties, keeping them informed, involved, and aligned throughout the software procurement process.

## 2. Will it scale and stay future-proof?

You may be making a software purchase to solve a specific problem you have now, but it's important to think bigger. Software that solves a single use case, be it predictive analytics for gas turbines, or production optimization for certain wells in a field, is an easy fix.

But software that scales to include more turbines or even hundreds of wells across multiple fields is another thing entirely. During your purchasing process, it's important to look ahead to emerging needs, even a year or two down the road, and ask yourself: can this software meet them?

## 3. Can I see a demo?

When it comes to software, seeing really is believing. PowerPoint slides and talented salespeople can only tell you so much. A demo will show you. It's within your rights to request this kind of real-life proof, to get a taste of what that technology can do for you.

## 4. Is it fit for purpose?

The harsh reality is that no software available today will be able to fit your needs perfectly. Out-of-the-box software products may meet about 80 percent of your needs.

And while this may not cover all your requirements, it's important to remember that, with the right software partner, it does mean that 4/5 of your requirements will be met within six months.

This is a far better option than holding out and waiting for the "perfect" solution (that may never come), or building something completely bespoke in-house that you will need to maintain indefinitely. This requires smart adaptability and a willingness to

go for "good enough" when it comes to software. Over time, and most likely faster than you think, you can go from good enough to game-changing.

## 5. How do you price it?

Pricing software is not copy-and-paste from the pricing of a pump. It requires an entirely different approach. Software is a product, and it is priced as such, rather than charging by man-hours based on level of experience.

Once you know the product price, ask about scalability, as future needs are bound to arise. Discuss with your vendor a pricing model that removes friction at later stages, for the time when you are ready to scale solutions across the rest of the organization.

For a complete list of essential questions and criteria, see the appendix—Industrial DataOps RFP Guideline: Questions to Aid Evaluation.



**Figure 17: Industrial DataOps Key Features Overview**



7:

# Quality Is King

Speed matters, but running fragile solutions leading to failures in production simply does not work in industry.



In 2014, even Facebook changed its motto for developers from "Move Fast and Break Things" to "Move Fast With Stable Infra". They feared that they may have been moving too fast to see where they were going clearly. As Mark Zuckerberg said, "What we realized over time is that it wasn't helping us to move faster because we had to slow down to fix these bugs and it wasn't improving our speed."<sup>21</sup>

Such caution may have been novel in the world of digital technology giants, but for the world of industry it is a well established necessity. Across heavy-asset industries, a whole range of weighty factors mean that the 'fail fast' mantra is rarely appropriate.

These factors include the risk of high-profile failures and outages causing serious public impact and lasting reputational damage, the cost of machinery and equipment, employee health and safety, and potential environmental impacts.

**"We can't operationalize unless we trust the data. If something fails and we can't provide auditability we are finished as an industry."**

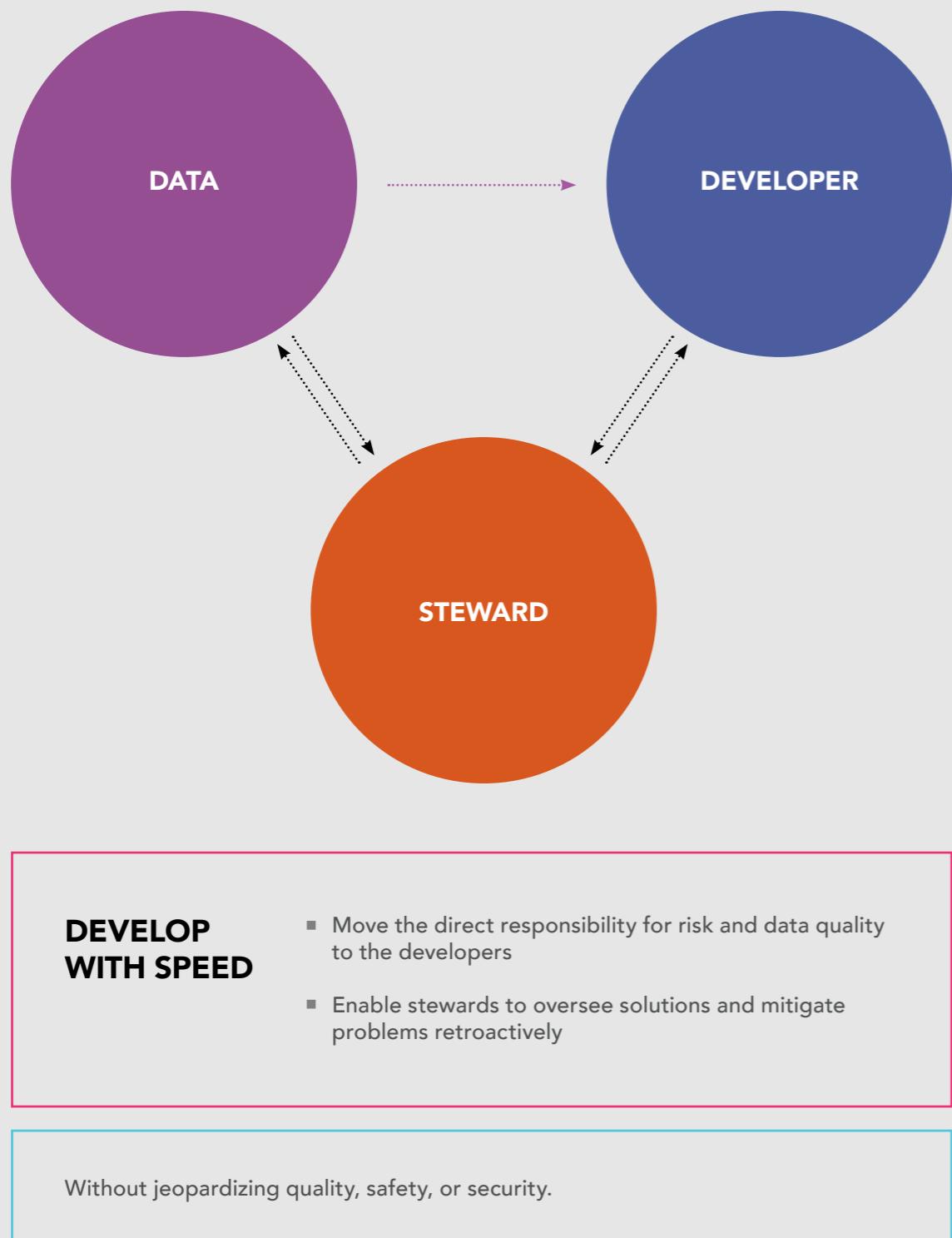
Stig Sund, Digital Operations Manager,  
Aker BP

For most of our software development and data engineering era (a relatively short one when compared to the industrial organizations being served), speed and quality have been in opposition. This is especially true when working with complex data dependencies in a real-time environment. Dependable, secure, observable data pipelines across batch and streaming data sources, as well as pragmatic data quality fitting the use case being developed, have long been aspirational but not achievable—at least not without majorly sacrificing speed.

With Industrial DataOps, organizations can achieve both quality and speed by moving the direct responsibility of risk and data quality to the development teams who understand the solution context best. This approach enables ambient data governance rather than slow, rigid, and costly centralized master data management.



Figure 18: Changing Through Tooling



"

**"The diverse and distributed nature of IoT solutions means that a traditional, one-size-fits-all, control-oriented approach to governance is insufficient. Organizations need to be able to apply different styles of governance for different types of data and analytics."<sup>22</sup>**

Industrial DataOps requires that organizations make use-case-relevant data quality controls easy for development teams consisting of business technologists, rather than data quality specialists.

This means seeking Industrial DataOps tools that provide most common data quality models pre-built and easy to apply in the natural application data template definition phase directly

22. Friedman, T & Judah, S (2020). Successful Internet of Things Initiatives Require Adaptive Data and Analytics Governance. [Online]. Gartner. [29 Mar 2021]. Available from: <https://www.gartner.com/en/documents/3986220/successful-internet-of-things-initiatives-require-adapti>

---

## Chapter 4

# Industrial DataOps in Action





Having covered the theory and key principles of Industrial DataOps, it's time to see how organizations are putting these into action in the real world, to achieve tangible improvements in the industrial value chain. This chapter explores stories from the field, including examples from oil and gas, power and utilities, and manufacturing.

The complex machinery and supply chains necessary to deliver the desired outcomes make general industrial value chains fall into the category of problems described as "systems of systems". In these, overall challenges are affected both by the complexity of individual processes and tasks, and by their interdependence.

Overcoming these challenges successfully calls for the combination of deep niche software (such as the control system of an individual oil or gas pump) with high-level software that orchestrates and governs enterprise-wide processes (such as an enterprise asset management system). Effective Industrial DataOps is imperative to assure the development, scaling, and managing of data-driven improvements that cross the intersections of different applications.

**Effective Industrial DataOps is imperative to assure the development, scaling, and managing of data-driven improvements that cross the intersections of different applications.**

One example would be developing a model to predict the optimal maintenance for an individual pump, where its output should trigger automatic creation of relevant tasks in the enterprise asset management system (EAM), considering supply chain, resource, and risk interdependencies. Industry is full of such potential improvements that can be made by linking relevant, contextualized data with real-time operations.

Taken in aggregate, such fusions of technologies, data, and automated processes are central to the concept of Industry 4.0, or the Fourth Industrial Revolution.<sup>23</sup> At the core of the necessary infrastructure to handle these improvements at scale lies industrial DataOps.

# Opportunities for Industrial DataOps

Let us first briefly discuss where the opportunities are. Industry's core driver is production, be that the extraction and processing of oil and gas, the generation and delivery of power, or the manufacturing of goods. Optimization of throughput is crucial, both in terms of production volume and, when the output is not a raw material, production quality. Your assets and equipment are vital in this, so naturally the maintenance of those assets is one of the main levers to increase throughput and reduce costs. The chief objective of a maintenance program is therefore to maximize uptime and minimize cost in an asset life cycle perspective. Production optimization and maintenance are both key areas ripe for data-driven improvement with an Industrial DataOps approach.

Managing field-worker efficiency is another key area of opportunity for improvement with Industrial DataOps. Often, practitioners separate equipment maintenance and asset integrity maintenance, where there will be different experts assuring the proper prioritization, risk assessment, and planning of the two.

The objective remains the same: to minimize negative production impact at lowest lifetime cost. The ultimate manifestation of the process is that some task is performed in the field by personnel, whether that task is a change of equipment oil, or the corrosion treatment of a handrail.

As well as production optimization, maintenance, and field worker efficiency, supply chain is also on the list of areas where Industrial DataOps can be transformative. We can add capital project planning and execution (in certain industries), as well as subsurface oil and gas with its exploration, drilling and wells, and reservoir domains.

The promise of Industrial DataOps is to increase the pace of deployment of data-driven improvements, including the scaling of any individual improvement across the fleet of assets, whether that asset represents equipment or larger facilities. Now let's take a closer look at some real-world examples of data-enabled use cases and the data operations necessary to support them.

Industrial DataOps in Action: **Examples from the Field**

# Extending the lifetime of critical assets and reducing service time

"

Up to 30% reduction in service costs





## Sector

Manufacturing

## Company

Aarbakke

## Challenge

Like many manufacturers, Aarbakke was facing performance challenges caused by underperforming critical equipment: specifically, computer numerical control (CNC) machines.<sup>24</sup> As a supplier to the oil and gas industry, Aarbakke must meet strict quality requirements compounded with short timelines.

Previously, service managers had depended on operators to log critical issues through emails or notes. Service managers would then physically go to each individual machine and manually review the local log to see the alarms. This legacy work process was limiting response time and resulting in lower throughput and poor quality events.

Aarbakke needed a new work practice to understand their critical equipment and provide sitewide visibility for active machine alarms, historical alarms, and near real-time sensor data. The team knew that using Industrial DataOps to increase visibility would enable them to better prioritize work, adopt more efficient work processes, and gain learnings to prevent repeat poor quality events.

## Approach

Aarbakke's chosen Industrial DataOps solution was used to integrate data from the manufacturer's source systems, contextualizing it all into a unified data model which was made accessible to all users. The data sources included both process (time series) data and events data (equipment alarms and maintenance activities).

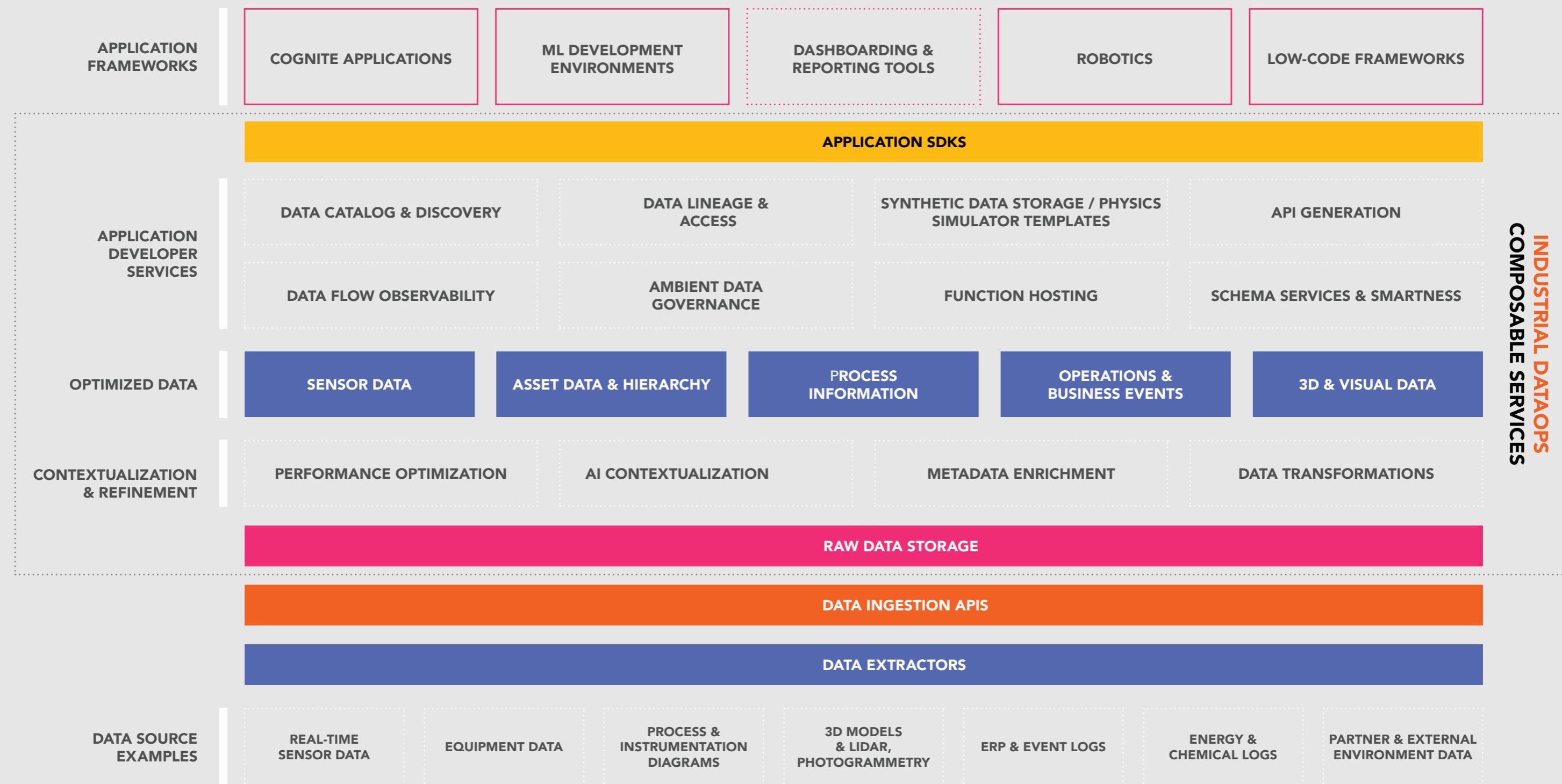
With a unified data model incorporating all the necessary data, it became possible to quickly develop an application to make Aarbakke's management of critical equipment more efficient. The delivered application provided a contextualized, site-wide view of all CNC machines that grouped alarms and events by assets. Service engineers could then use persona-based filters to take targeted maintenance actions. Insights created from this application reduced service cost by 20-30% and delivered both improved quality and increased throughput.

## Impact

- 20-30% lower service cost by extending the lifetime of the machines, with a reduction in the number of breakdowns and overall downtime.
- Faster response times to issues with organizational visibility.
- Improved quality and increased throughput.
- Prioritization of assets to intervene before unexpected failures.
- Better support for maintenance workers in the field.



**Figure 20: An Industrial DataOps Platform Services Architecture**



## Industrial DataOps in Action: Examples from the Field

# Supporting maintenance workers out in the field

**Up to 55% increased worker productivity**

### Sector

Manufacturing

### Company

Yokogawa

## Challenge

Large-scale digitalization of the manufacturing industry will require all relevant data to be accessible by users out in the field.<sup>25</sup> However, for many manufacturing companies, data is trapped in complex, siloed systems without the ability for maintenance personnel to easily access the information they need when diagnosing equipment, performing repairs, or completing inspections. Siloed data creates inefficient work processes for maintenance personnel in their day-to-day activities, often requiring them to access multiple systems to find the information they need.

## Approach

In less than two days, the right Industrial DataOps solution liberated and contextualized data from Yokogawa's source systems, including process variables, equipment information, historical events, and instruction manuals. This short turnaround was achieved using contextualization services to automatically create relationships between process variables, equipment, and events for plant assets.

This contextualized data could then be connected to a field application designed to meet the needs of the digital worker. Accessible from tablets and mobile devices, this made real-time process data, historical data, documentation, CMMS work orders, and pictures available to maintenance personnel out in the field. Assets can be identified simply by scanning the tag on any piece of equipment to see all relevant information.

Having this data at the tips of their fingers significantly increased the efficiency of maintenance personnel and equipped them with the information needed to perform their daily tasks.

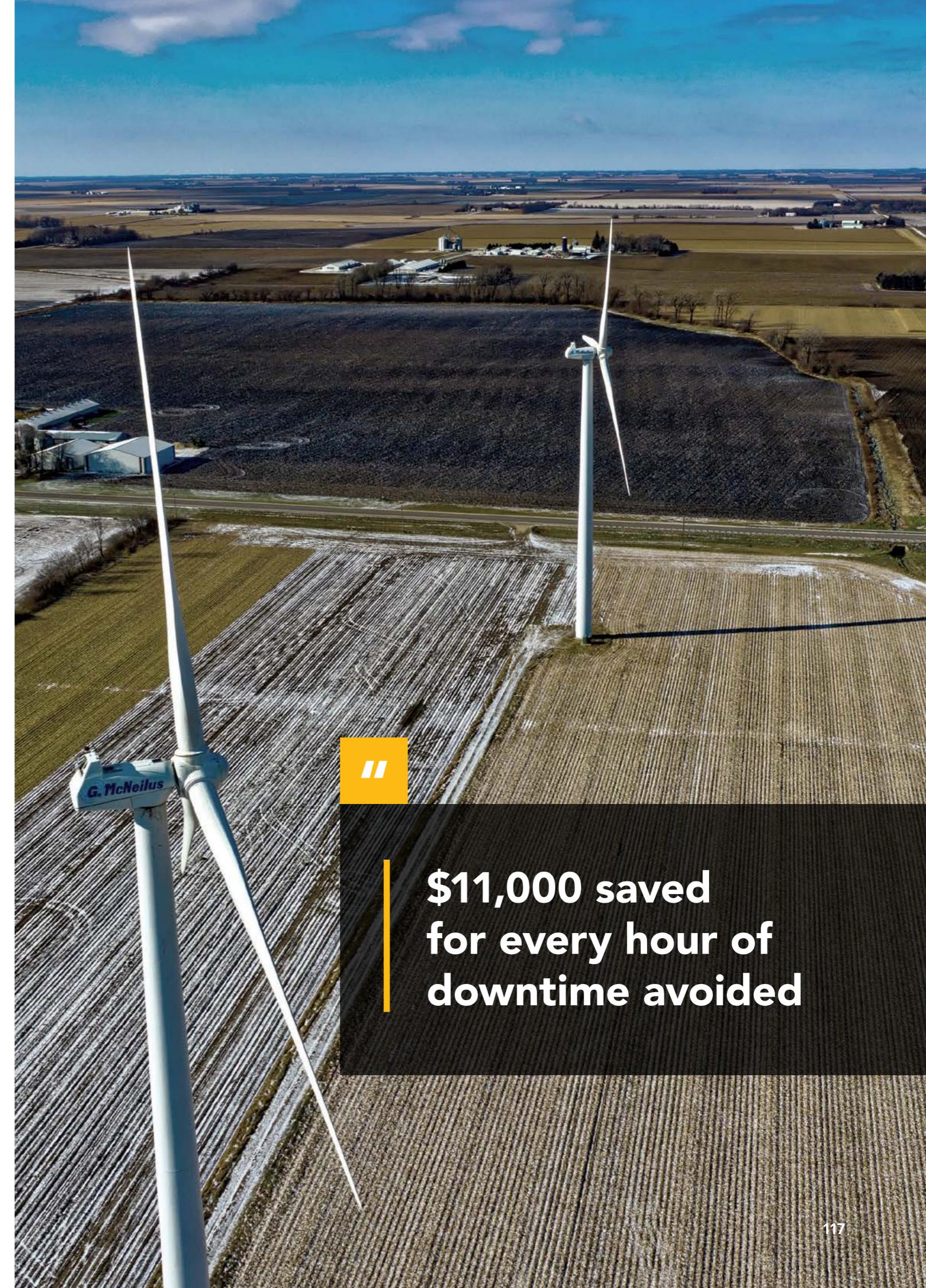
In addition, an operational digital twin was created, combining the liberated, contextualized data with a 3D model. This was built in under one hour, after capturing some 400 pictures of the Kofu plant. The contextualized real-time process and historical data was then overlaid in the 3D model, giving users a powerful visualization tool to explore the plant. These 3D models can be used by maintenance personnel to better understand how a piece of equipment fits into the overall process and see other work orders created in the same area—all from their mobile devices.

## Impact

- Maintenance personnel are able to access all of the relevant information they need when completing work out in the field. With mobile access to all the process and asset data, they can quickly diagnose errors and conduct maintenance work more efficiently, resulting in 30–55% increased worker productivity.
- The 3D model with contextualized data saves field workers time with more efficient processes for locating equipment and planning work to be completed in the field, improving off-site planning and support.
- The solution delivery time is less than one week, delivering value within weeks of implementation.

Industrial DataOps in Action: **Examples from the Field**

# Optimizing turbine start-up sequences to save costs



\$11,000 saved  
for every hour of  
downtime avoided

**Sector**

Power and utilities

**Company**

Hafslund E-CO

## Challenge

The demand for power fluctuates during the course of a normal day. It bottoms out during the early morning hours, then begins to climb as people wake up and go about their day.<sup>26</sup> After a brief dip following lunch, the demand reaches its highest point in the late afternoon as people return home, before it slowly tapers off as the evening turns into night.

Grid operators need to find the power sources to cover this demand on a daily basis. Scandinavian utilities, for example, trade power on the regional exchange market Nord Pool. The existence of multiple power generation sources such as hydro, wind, coal, and nuclear, presents a challenge for all the players. In order to ensure that demand for electricity is met at any given time of the day, utility companies are often required to suddenly ramp up or ramp down production. This presents a major opportunity to companies that can efficiently stop and start their production, such as Hafslund E-CO, Norway's second largest power generation company.

At many hydroelectric power stations, including Hafslund E-CO's in Nes, Norway, the start-stop sequence consists of about 10 steps, which may include checking auxiliary systems, different opening and closing processes, and breaker and generator procedures.

Different alarms may appear during each step of the sequence, but it is not always clear to the engineers which step triggered an alarm. Since an alarm may stop the entire sequence and prevent the turbine from starting, more information would help the power station engineers optimize the process and avoid revenue loss.

## Approach

Hafslund E-CO worked with developers at their new Industrial DataOps provider to enable live streaming of sensor values and alarms from the utility company's control system into a data fusion system. Using an application for production optimization, the developers, together with Hafslund E-CO's domain experts, created a dashboard that tracks the duration of each step of the start-stop sequence, displays relevant sensor values, and categorizes unsuccessful attempts.

The dashboard notifies Hafslund E-CO's personnel immediately about alarms and lets them quickly check how the plant is performing at any time. It also enables closer monitoring and analysis. For example, the dashboard may log that the typical opening time for the main inlet valve is 10 seconds. Over time, if that subprocess begins to take longer, the dashboard will flag the issue, suggesting that the valve system should be scheduled for maintenance.



## Impact

- Live streaming of sensor values and alarms from the utility company's control system makes it easier for engineers to troubleshoot alarms faster and take action.
- Dashboard enables monitoring and analysis, helping engineers troubleshoot alarms and improve uptime.
- Hafslund E-CO's engineers have a greater ability to stop and start production efficiently.
- For every hour that Hafslund E-CO avoids downtime at the Nes power plant, which has a production capacity of 250 MW, they also avoid revenue losses to the tune of \$11,000.



<sup>26</sup>. Cognite. c2020. How Cognite Helped Hafslund E-CO Optimize Turbine Start-up Sequences. [Online]. [17 Mar 2021]. Available from: <https://www.cognite.com/customers/dataops-power-utilities-turbine-startup-optimization>

## Industrial DataOps in Action: Examples from the Field

# Enabling data-driven slug prediction to boost production

**1% increase in production**

### Sector

Oil and gas

### Company

Aker BP

### Challenge

'Slugging' is a common production challenge in the oil and gas industry. This means a separation of the three-phase flow (gas, oil, water) in the pipeline, and the accumulation of one or more of the phases blocking the flow.<sup>27</sup> The factors that influence slugging can be either transient (such as the opening or re-routing of wells) or steady-state. Therefore, simulators and real-time production data need to be combined to monitor and prevent slugging.

### Approach

Aker BP's Industrial DataOps solution enables access to thousands of live and historical time series, continuously analyzing these for pattern recognition and statistics, ensuring the necessary data quality for operational decision-making. The live production data was integrated with a data-enabled third-party application to deliver the complex simulations and models required for slug prediction.

Live operational data was fed into hybrid models with self-learning algorithms that could identify field behavior and generate predictive models to identify slugging scenarios. Aker BP was able to develop optimization models which deliver real-time, actionable insight to production engineers, to avoid slugging incidents and the associated production losses.

### Impact

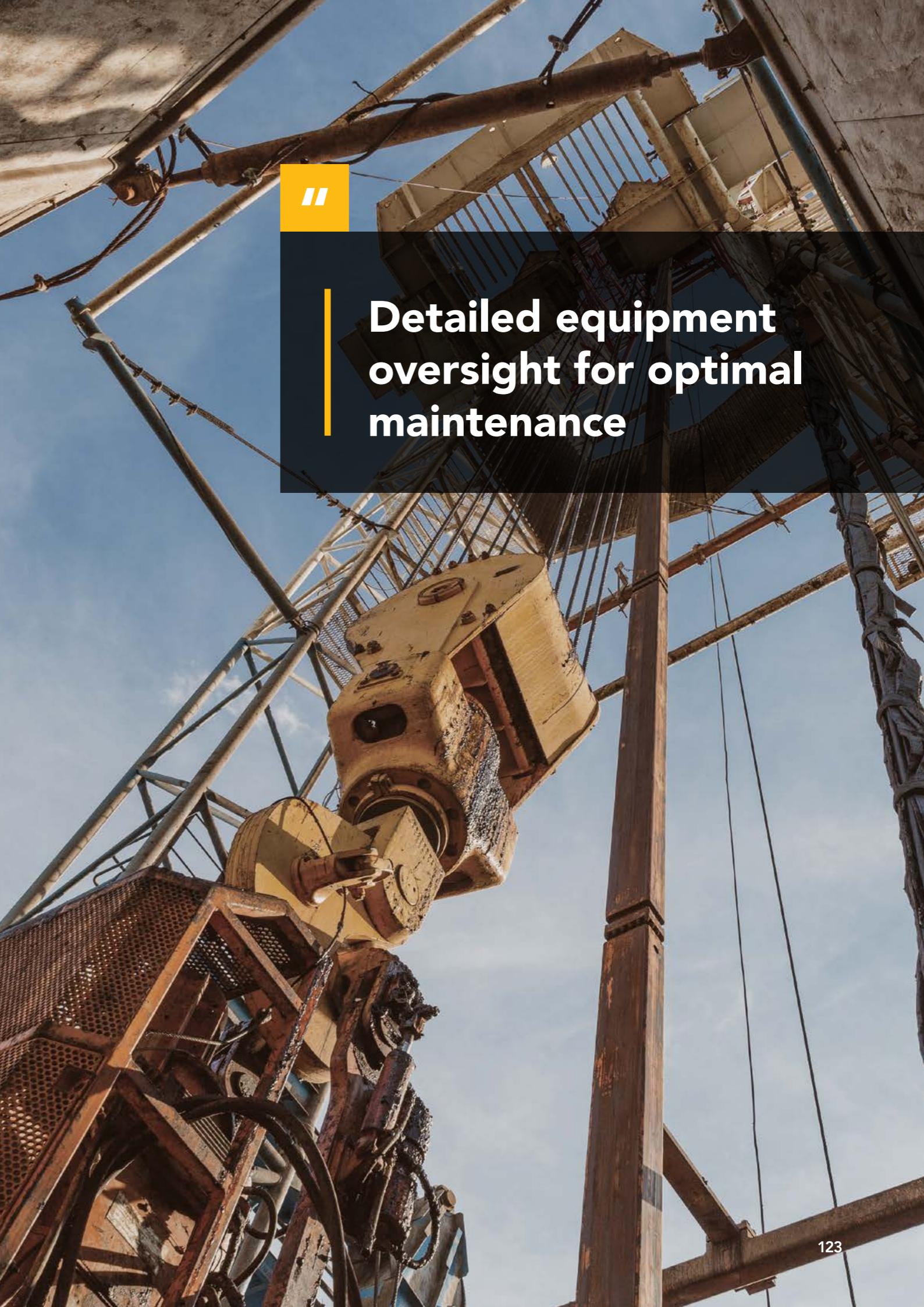
- The company saw a 1% increase in production.
- Slug handling and prediction capabilities improved significantly.
- Production engineers gained real-time, actionable insight.
- Operators received user-friendly decision support and algorithms that deliver early warning signs of imminent slugging.

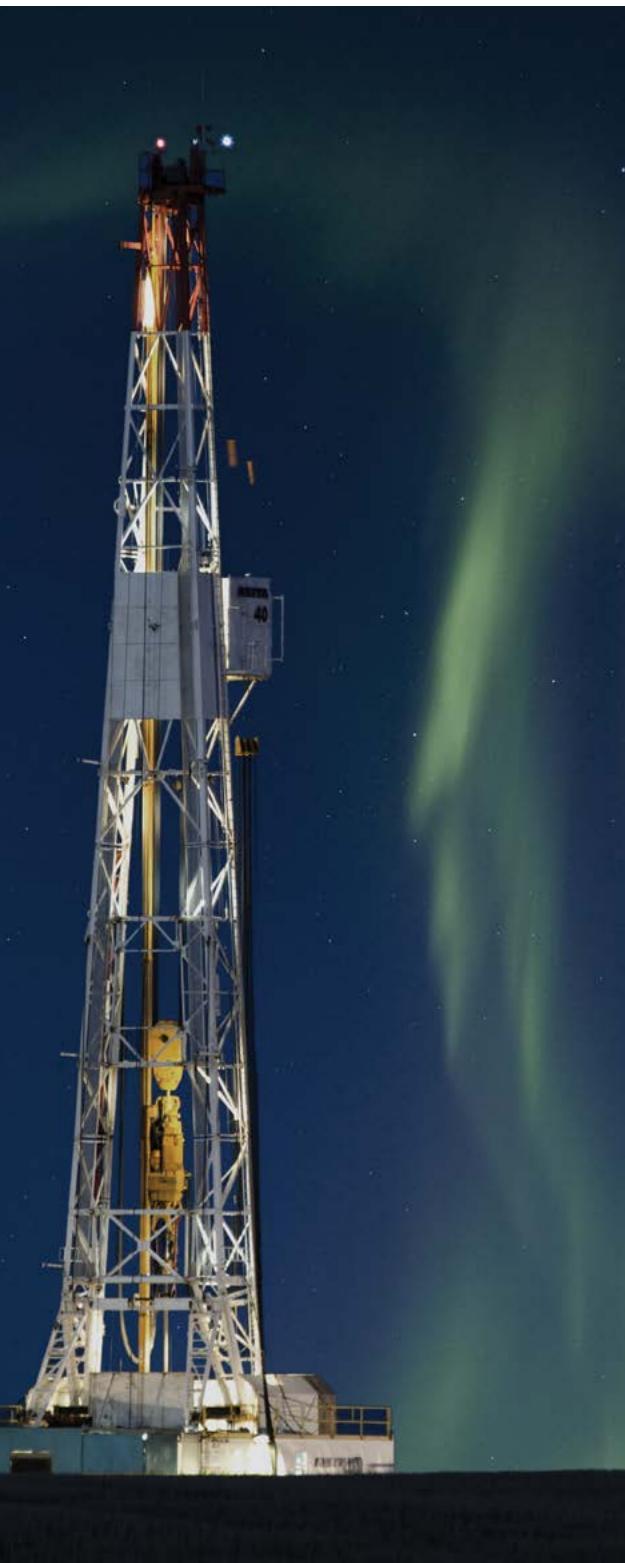


Industrial DataOps in Action: **Examples from the Field**

---

# Enabling data-driven maintenance and performance- based service delivery





## Sector

Drilling and wells

## Company

MHWirth

## Challenge

Leading drilling technology and service provider MHWirth landed a new performance-based maintenance contract with a strategic customer, for surveilling and maintaining their fleet of drilling equipment on rigs across the world.<sup>28</sup> To meet their contractual commitments, and keep the account profitable, MHWirth needed to rethink how they utilized data in their maintenance decision-making. The need for taking a condition monitoring and predictive maintenance approach was a given, but there the challenges began. How could they set up their IT architecture to efficiently survey and detect anomalies from a worldwide fleet of assets? It was also crucial to ensure the necessary data governance and quality to make actionable decisions, which could potentially make or break the customer relationship.

## Approach

MHWirth implemented an Industrial DataOps solution to capture and organize live equipment data. The provider set up live extractors of the necessary data, creating asset templates and the supporting data models as a foundational piece of the underlying data operations.

With all the relevant data in the system, MHWirth could use the visualization and analytics tools of their choice to create digital representations of the current state of their drilling equipment. The tools enabled MHWirth to perform predictive analysis on the data to better plan the right maintenance program. The stream of real-time data fed directly into the model, then the results were made available for other visualization tools, applications and machine learning models.

MHWirth dashboards now use both historical and real-time data to inform experts about the true condition of equipment. This is the prerequisite for assuring maximum uptime of drilling equipment in operation, balancing appropriate maintenance with a profitable maintenance contract.

## Impact

- By bringing all industrial data together in one place and linking it automatically to an asset hierarchy, MHWirth gained a better understanding and fuller control over their industrial reality.
- Experts analyzing the data in MHWirth's system can rapidly identify which equipment needs service, prioritize their actions, and advise on optimal maintenance.
- The company has now the opportunity to develop truly insightful maintenance programs, keeping maintenance costs in check, extending the lifespan of equipment, improving equipment reliability, and minimizing unplanned maintenance and downtime.
- MHWirth now has continuously innovate on the services they provide to drilling companies, bringing condition-based maintenance services into the mix.



## Industrial DataOps in Action: Examples from the Field

# Scaling operational impact from visualization software

**Complex, large-scale data sets visualized to aid better decision-making**

### Sector

Oil and gas, manufacturing, power and utilities

### Company

Various

### Challenge

Visualization software, like Power BI, Tableau and TIBCO Spotfire, is commonplace for all companies attempting to become more data-driven.<sup>29</sup> However, visualization software is nothing without data. Unless the data can be found, insights cannot be generated; and unless the data can be trusted, the insights are not actionable. These problems are well known in large organizations with many different data domains, data stores, and data owners. Getting IT to provision the data can be cumbersome in itself, but just as often, the user does not even have the full picture of what data exists. When the data is found, visualized in Power BI and used to generate insight, constant assurance of data quality is necessary to create sufficient trust to utilize the data in operational decisions. All these challenges need to be overcome to scale the operational impact of visualization software.

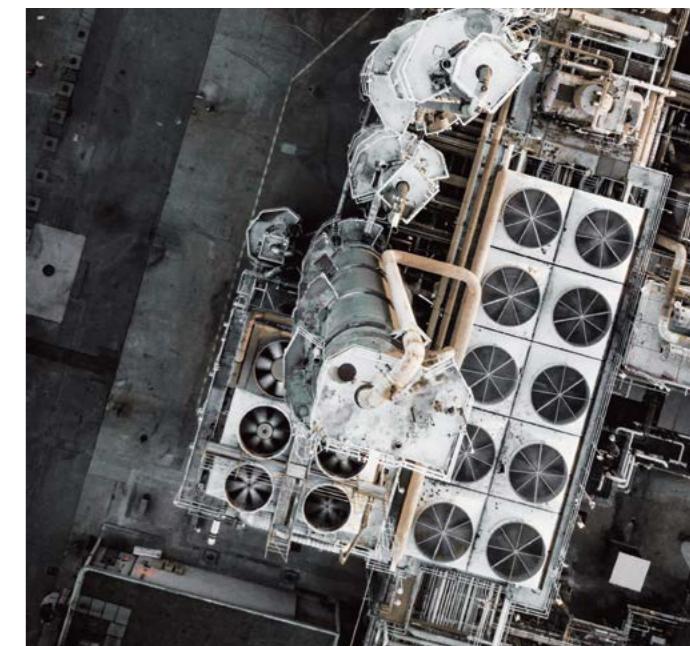
### Approach

Using the right Industrial DataOps solution, business users are easily browsing all available data and provisioning it to their visualization software of choice. By also providing data management and data quality tooling, users easily set up data lineage and quality monitoring. This ensures the visualization is actionable at any given time, and ripe for operational decision-making.

As the data is provisioned through the solution, users are building visualizations independent of the operational and IT systems where the data resides. Integration with AD providers enables easy sharing or restricting of data access to visualizations. Enterprises are able to become more data-driven without the need for complex infrastructure projects on the IT-architecture side and long, expensive training programs on the user side.

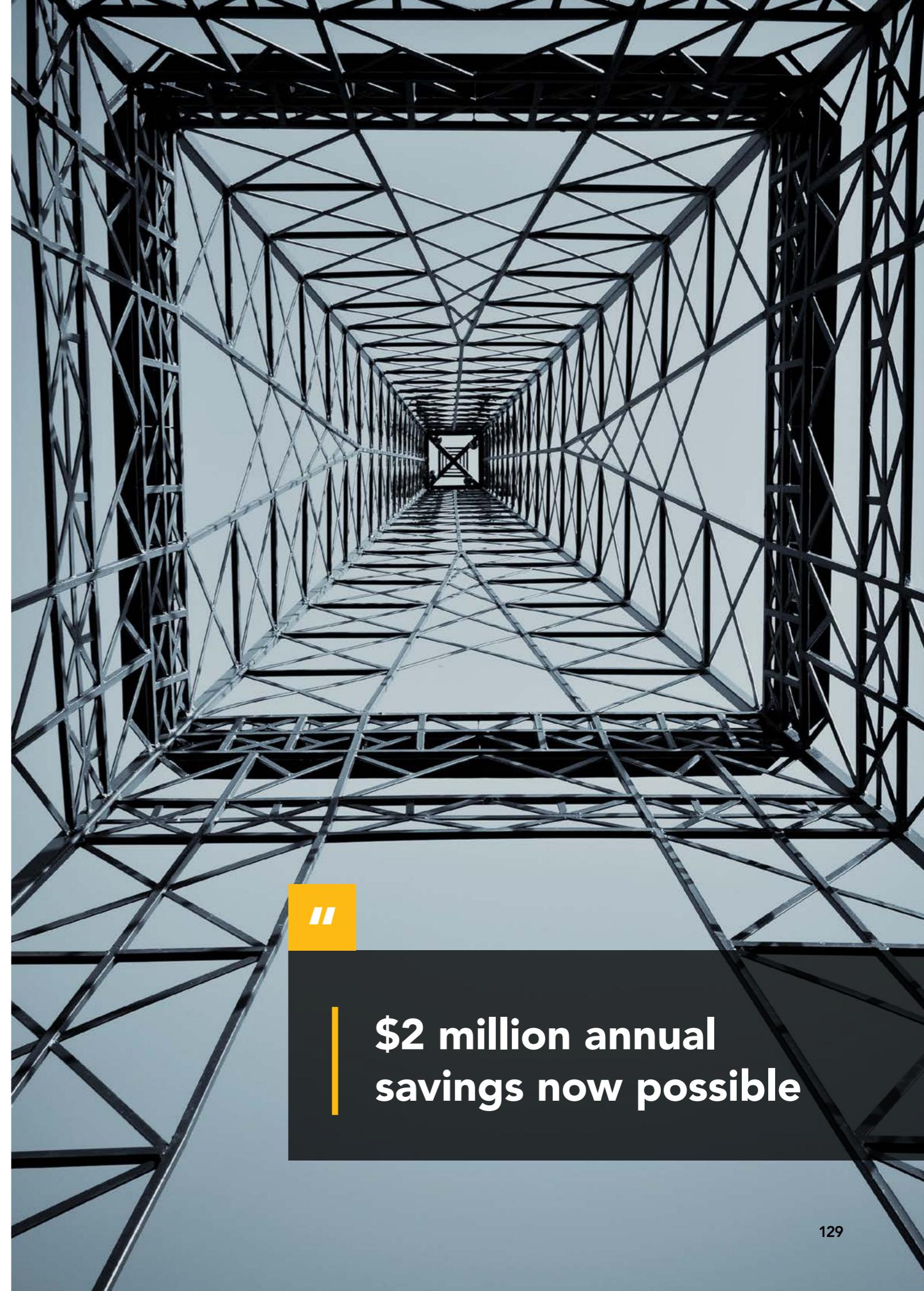
### Impact

- Organizations gain access to visualization of complex, large-scale data sets.
- Large-scale data visualization is available to all necessary users in the organization, in a simple browser view.
- Users gain instant access to use cases, to fuel innovation.



Industrial DataOps in Action: **Examples from the Field**

# Preventing expensive transformer failures for an energy grid operator



"

\$2 million annual  
savings now possible

## Sector

Power and utilities

## Challenge

Transformers are some of the most expensive and critical components in a power grid.<sup>30</sup> Often weighing in at more than 200 metric tons, these massive devices are situated at critical points in the grid, transferring electricity between alternating-current circuits and increasing or decreasing the voltage as necessary.

Grid operators sometimes experience transformer failure. These events can lead to power outages for consumers and production losses for power companies. In the worst-case scenario, a malfunctioning transformer can catch fire and even explode. Both repair and replacement are expensive and time-consuming.

A major grid operator, responsible for hundreds of transformers, was experiencing about one malfunction a year. The grid operator was conducting regular maintenance of the transformers, and it also invested in replacement components to ensure that power could quickly be restored in the case of an outage.

The grid operator expressed an interest in improving both of these processes. How could data help the organization identify early signs of transformer failure, and how could it optimize its spending on replacement parts?

## Company

Major grid operator

## Approach

The grid operator worked with an Industrial DataOps provider to liberate information about transformers from its source systems, including temperature, load, dissolved gas analysis, technical specifications, and inspection logs, and ingest it into a data fusion platform.

With access to all the data relevant to transformers in a single location, the development team was able to calculate a health index for every transformer in the power grid. That health index was then visualized in a dashboard, giving the grid operator's engineers the ability to monitor the entire fleet of transformers at a glance and see which components should be prioritized for maintenance.

## Impact

- \$2 million savings per year driven by failure reduction of 20-50%.
- The health index helps the grid operator make data-driven decisions about how to plan its transformer maintenance activities.
- Each transformer failure costs the grid operator at least \$5 million. The grid operator has set a goal of reducing the chance of failures by 20-50% over the next five years, which in the short term will save the company about \$2 million a year.



---

## Conclusion

# Industrial DataOps Is the Future



**"On one hand, change is simple—more of the same. The problem is that 'the same' now means everything will continue to change, just faster."**

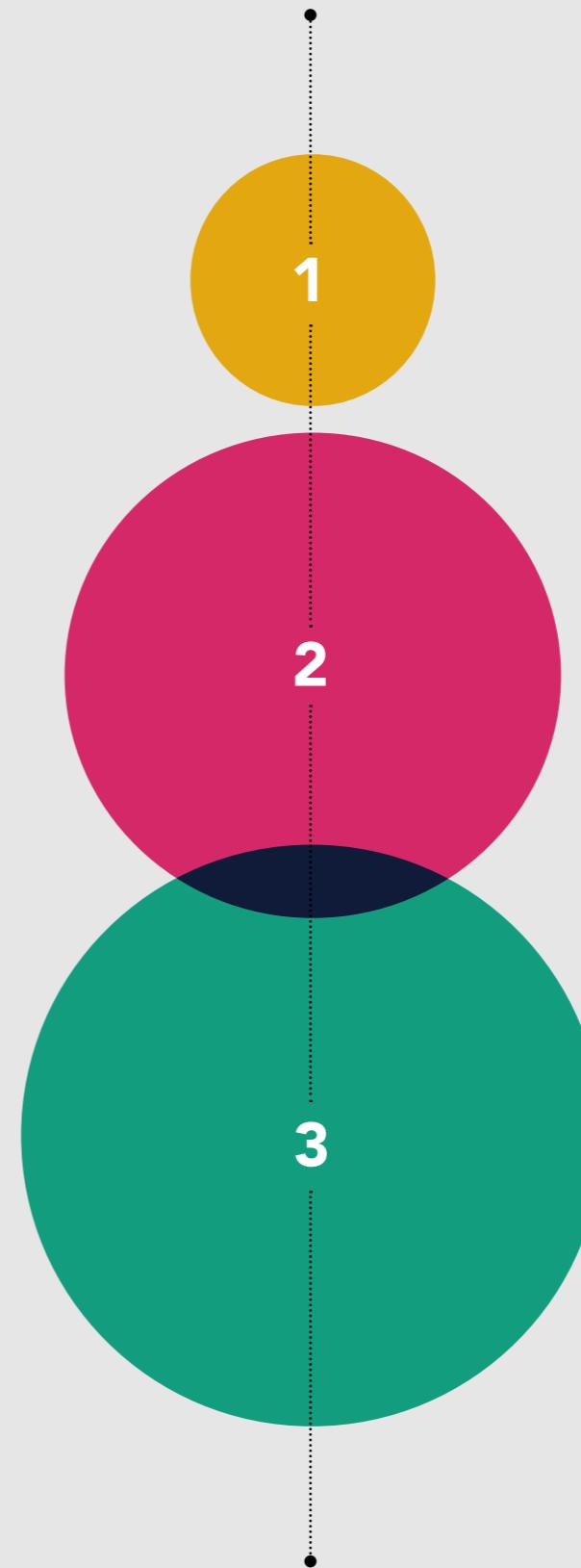
William Mayo, CIO, Broad Institute of MIT and Harvard.<sup>31</sup>

As we hope this book has made clear, the technology and know-how to develop, refine, and scale the use of Industrial DataOps are falling into place. Across industries, DataOps is ever more accessible through nascent platforms under rapid development and refinement by industrial technology and software developers. It's also becoming better understood.

Industrial DataOps is the single most promising path for digital, sustainable, financially viable transformation of asset-intensive industries. Much of the promise of this discipline is that it depends very much on democratizing innovation. It needs the power and creativity, en masse, of both the existing and future industrial workforce.

It assumes that more widespread, equitable, and likely faster digitalization of industries will come from making data speak "human," rather than placing the world's trust in a small fraction of humans speaking "data" (that tiny percentage of people with the right training and brain wiring).

Delaying the awareness-building, consideration, and adoption of Industrial DataOps, in our view, puts companies, even entire industries, at risk of losing out on once-in-100-year opportunities. These tend to emerge from periods of seismic change, like the one we enter together now.



**Industrial DataOps, or data operations for industry, is the clear frontrunner to become the driving force for transformation in industry.**

This happens first by **making your data available**, by identifying where it is, how to get at it—and how to store it for later use.

The next step is then **making your data useful**—freeing it from silos and making it speak "human", so that its value can be released across operations. Contextualize it and design homes for it such that all of the minds and functions in your organization can actually understand it, use it, and innovate on top of it.

And finally, **making data valuable**. Extracting the maximum value depends on being able to obtain insights that inform better decisions, and enabling all business users to become solution creators. Scaling these benefits brings data value to the whole organization.

# Getting Started

---

Industrial DataOps admittedly requires you to do some prep work. You need agreement and unity among all data stakeholders in your organization. As we've said, that kind of cohesion and cooperation needs to go beyond just making data available and allocating resources for that. Your business functions will need to be geared and organized for close involvement in data projects well beyond resource allocation, proof-of-concept, and scaling phases.

As we've said, this is not a quick and easy place to get to. It requires a longer-term view and committed leadership to bridge divisions and align your vision. But the commitment will pay ample dividends and shore up your future success.

# Are You Ready?

---

Getting through this book means there's a chance you want to be. Assessing your readiness for Industrial DataOps will need you to assess your organization's digital maturity first. As we said in Chapter 1, digital maturity is a key metric of digital success. This also bears repeating: getting real ROI from digital initiatives is more likely to come from steady innovation and long-term strategy, than from a quarter-long burst of digital enthusiasm to impress stakeholders (or shareholders).

A sign of digital maturity is that you're able to naturally focus on the long game, building through tools and processes so that digital becomes effortless across a broad range of stakeholders. Another sign is that you're able to measure, measure, and measure—at both the granular and the holistic levels.

We'll say it again: digital maturity spans people, processes, and data. It is very much a multidimensional means of change management.

**Evaluate on Purpose & Strategy, People & Mindset, Value Capture Process, Information Architecture, and Data Ubiquity, and go from there.**

# What Industrial DataOps Can Do for You

As we saw in Chapter 4 (Industrial DataOps in Action), some leading organizations have made an early start and are already leveraging Industrial DataOps to AI, data-intensive applications, complex research, and analyses.

Some other great examples come via researchers at MIT, who interviewed a number of companies to determine how they were capitalizing on the new approach. The results revealed a range of highly beneficial use cases from industry.<sup>32</sup>

One of the organizations is “processing and analyzing an extraordinary amount of data using modern data management practices to help automate shipping fleet maintenance”. To do this the company is using predictive analytics on “trillions of data points” to estimate repairs and breakdowns.

Another company is focusing on monetizing data to help reduce shipping fleet downtime, which saves time and costs while keeping the fleet operational. The organization identifies trends toward failure before they become alerts, and utilizes the data to understand present and future performance of the vessel.

Autonomous vehicle operation is another area poised to benefit from Industrial DataOps. As major auto companies double down on R&D in autonomous driving, enormous sums of data, including telemetry and imaging, will be generated by autonomous fleets and hardware (traffic lights, road sensors) to further leverage as a basis for innovation.

**On a single asset alone, Aker BP is recording annual savings of \$6 million per year in saved time and thwarted production losses.**

In the world of oil and gas, Norwegian operator Aker BP is deploying an Industrial DataOps framework in its operations to comply with regulations on oil in water and reduce production losses. In essence, they have implemented a DataOps-powered smart monitoring system that visualizes all data relevant for troubleshooting water contamination, and a recommender system with an underlying machine learning model to identify the worst actors related to high oil-in-water concentrations.

The smart monitoring system displays near real-time data from their Industrial DataOps platform and visualizes it in an intuitive dashboard. Additionally, calculations combining sensor values and simulator outputs provide engineers with virtual sensors and physical properties they otherwise would not have had readily available.

On a single asset alone, Aker BP is recording annual savings of \$6 million per year in time and averted production losses, all while protecting the local environment and complying with ever tighter environmental regulations.



Imagine this single case, focused entirely on detecting water contamination, scaled up over an entire set of assets or an entire field, across operators. The potential for colossal, positive transformation is evident.

The effects are even more powerful when data sharing is made possible. Aker BP works with Framo, a major supplier of submerged cargo pumps to the oil and gas industry. Their Industrial DataOps platform enabled the secure sharing of selected live data between their two organizations. Framo used Industrial DataOps to access Aker BP's industrial data, which helped inform their product development, leading to a more sustainable, performance-based business model. The solution cut emissions and waste by reducing maintenance needs by 30 percent and shutdowns by 70 percent, and by increasing pump availability by 40 percent.



## The potential for colossal, positive transformation is evident.

The seeds have been planted for solutions like these to scale massively over the coming years. Ever-growing sources of data will intensify the need for smart extraction, liberation, contextualization, and analysis, underlining the need for Industrial DataOps.

Companies who commit to this will reap the benefits of converting meaningful data into impactful, measurable action. Meanwhile, the sophistication of users will develop and improve, and tools that were once overly technical and confusing will become powerful platforms of change, even for the digital layperson.

The ascent of Industrial DataOps is now bringing on a period of unprecedented collaboration between humans and machines. This potential to essentially synchronize human knowledge and creativity with advanced technology is still largely untapped in the industrial world. This is without a doubt one of the most significant big-picture opportunities of Industrial DataOps.

# Seize the Industrial DataOps Century

We see a dazzlingly bright future ahead for the organizations, irrespective of industry or market capitalization, who take on the brave new world of Industrial DataOps. Many of the companies (and industries) still catching up to the potential of the new approach are those who already have rich, heralded legacies, having irrefutably and permanently changed the world of the 20th century.

It's time now for organizations to adapt and build new legacies, not planted in the past, but projected forward, on a mission to change the world of the 21st century. For these companies and industries, Industrial DataOps will be their most valuable tool.

---

## Appendix

# Industrial DataOps RFP Guideline

### Questions to Aid Evaluation



So you have established that Industrial DataOps can play a vital role in your effort to truly transform your business. The challenge now is to define what capabilities your Industrial DataOps solution needs to support your business.

This section provides a guideline to build out your request for proposal (RFP) and ensure you account for all critical capabilities and functionalities required for success.

**As no single solution will solve all your data challenges, you will need to align your organization around the right capabilities critical to unlocking the potential of your industrial data.**

This guideline will present the key areas to consider and should be used as a starting point to build a framework tailored to the needs of your organization.

Issues to consider are presented in the form of questions. You may choose to put some or all of these directly to potential solution providers, as well as using them as an internal assessment tool.

# What to Consider When Creating a Request for Proposal (RFP) for an Industrial DataOps Solution

| Use Cases and Past Successes

| Functionality

| Solution Architecture

| Project Execution, Services, and Support

| Security

| Usability

| Software Maintenance

| Future Development

| Pricing Model

# Use Cases and Past Successes

First and foremost, Industrial DataOps must be able to deliver long-term value to your organization. Making this happen requires alignment between your organizational goals and the potential solution provider's capabilities. Knowing that your solution provider has competency within your domain will de-risk the probability of under-delivering on your expected ROI.

## Questions to evaluate a potential solution provider:

- Can you provide a brief description of your company, industrial business areas, main products/services, relevant expertise and business strategy?
- Are your products/services general or specific to the relevant industry? Can you describe your domain expertise?
- How would you describe your key product differentiation?
- What is your experience with helping clients build business cases and developing a target ROI? Can you provide examples of successful business cases delivered?

*Expert Tip: Successful industrial DataOps solutions should start with 1–2 use cases defined before any work begins. Have a backlog of 2–5 more to move on to once success is achieved with initial use cases.*
- How many existing customers do you have? Are there past successes you can share related to our industry?

- Does the proposed solution enable more effective asset management? Can you provide examples?
- How have you applied machine learning solutions to solve client use cases? Can you share any use cases utilizing hybrid AI solutions (combination of physics and ML capabilities)?
- What use cases have you delivered regarding unstructured data (e.g. video, 3D)?
- What are the most common types of use cases you have delivered?
- Do you have reference customers we can talk with?
- Can you provide a product demo?

# Functionality

Properly assessing Industrial DataOps software requires an understanding of two components: the foundation and the connectivity. Assessing the foundation is critical to ensure that the proposed solution will support your industrial data use cases and provide the tools needed to minimize time to value, and maximize scalability and repeatability.

Connectivity has two components: data extraction and application layer. Data extraction capabilities must allow you to connect to both existing and future data sources. The application layer focuses on how the solution provider will support applications on top of the foundation to deliver use cases.

## Questions to evaluate a potential solution: Foundation

- How does the solution perform data contextualization (data mapping)? Is it automatic or semi-automatic? Does the solution suggest relationships to make identification and construction easy?

*Expert Tip: The ideal solution should automate this process as much as possible, otherwise manually expanding the system to include new data sources will be extremely time-consuming and hard to manage.*
  - How is the contextualization (data mapping) process managed? Is it easily accessible?
  - How do users make edits?
  - How is the data model created in the proposed solution? How are relationships between data sources managed?
  - What types of data formats are supported in the proposed solution?
  - How does the proposed solution support data visualization?
- How does the proposed solution manage data quality? Are rules pre-built? Can rules be modified? Are rules applied universally or per use case?

*Expert Tip: Data models are designed to be reused. Data quality should have the flexibility to be applied per use case. For example, different use cases may require the same data, but using this data for remote monitoring of an asset will not require the same update rate as using the same data to run an analytics model measuring performance.*
  - Does the proposed solution support templatization? How can applied work be reused?

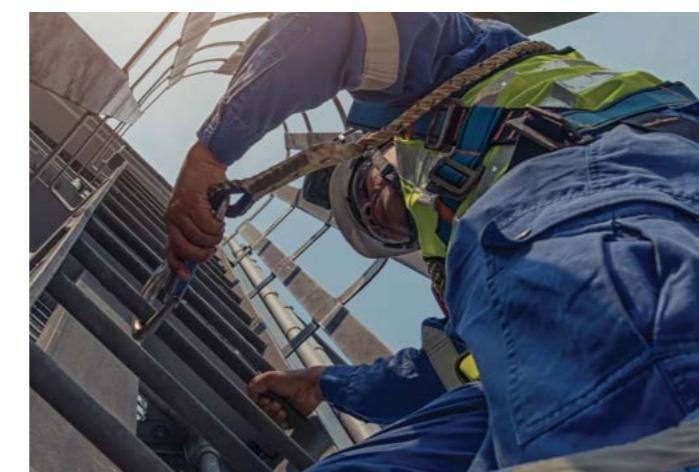
*Expert Tip: Templatization is a key component to scale solutions and ensures your organization will avoid getting trapped in Proof-of-Concept purgatory.*
  - How are notifications/messages supported in the proposed solution with regards to users associated with data and administrators? How does the solution score on scalability?

*Expert Tip: As you expand beyond initial use cases, you will need a solution that is scalable. Industrial DataOps should be able to address scale at both site and enterprise levels.*

- How does the solution support trending analysis of the data? How are trends visualized and reported?
- Can the solution analyze trends in data quality and predict when metrics will exceed predefined thresholds?
- How does the solution document completeness (integrity) of the ingested data and ensure data is not lost in transit?
- How do you work with third-party vendors? Which have you worked with in the past?
 

*Expert Tip:* Look at examples of proven solutions with third-party vendors so you can have confidence in being able to connect your disparate data sources.
- Is the front-end framework built on open standards? How do you support open front-end frameworks?
- How does the solution ensure that data is processed quickly and readily made available e.g. time series data?
 

*Expert Tip:* Access to centralized, remote, relevant time data creates opportunities for many new use cases at both the site and enterprise levels.
- Does the solution require plugins such as Office or Flash?
- Is it able to ingest both tabular and graph-structured data without loss of information?



## Connectivity

- When receiving asynchronous time series data, how does the solution handle time-stamping?
- Is the solution able to handle data inserts, updates, and deletes by itself?
- Does it support multiple modes of operation, such as batch and stream-based ingestion and in-memory versus persistent data storage?
- Does it follow agile development principles and how do you ensure it is up-to-date on market trends and technical standards?
- How does the solution support compression of data and metadata?
- Does it report the source for each data point, event, and time series, plus associated metadata for users to assess the data quality?
- How are the metadata fields of existing data and metadata updated? How are updates executed and managed?
- How is the connection between data and metadata made? Are they stored or linked? Can metadata be linked to several data entries?
- Expert Tip:* Prebuilt data extractors should exist for many open protocols and advanced Industrial DataOps solutions will have existing extractors to individual industrial solution providers such as Siemens, ABB, and Emerson.
- How easily can we (the client) develop our own applications on top of the product?
 

*Expert Tip:* Further assessment is needed when thinking about application development for data engineers and domain experts. Proposed solution providers should have pre-built connections to well adopted applications such as PowerBI or Grafana.
- Does the proposed solution provide an associated SDK? What languages are supported?
- What types of underlying data sources are supported? What connections are most common?
- What is the solution's capability of accessing real-time data? What are the scalability limitations to this capability?
- Does the solution have connectivity and native access to relational databases?
- Does it have connectivity and native access to non-relational structures?
- How do you ensure interfaces for data exchange (such as REST APIs) are kept stable and robust to changes?
- Does the solution support versioning for continuity so that both new and previous versions of data pipelines are supported? Can versions be rolled-back?
- Does the solution support a layered and scalable REST API?
- Is the REST API stateless, enabling easy caching with no need for server-side state synchronization logic?
- Can underlying data be exported from the proposed solution as a CSV or XLSX? Is data and metadata exported in standardized formats?
- Are there any limitations to the ability to extract historical data?

# Solution Architecture

Every organization will have unique architecture requirements that should be addressed from the beginning. The key here is to ensure that the proposed solution provider is set up to meet the requirements of your existing environment.

## Questions to evaluate a potential solution provider:

- What are the key components of the proposed solution and how do they operate/interconnect?

*Expert Tip: Any architectural requirements can be included here. Many organizations have already made investments to integrate OT/IT data silos into data lakes or data warehouse solutions. Your Industrial DataOps solution should leverage the investment into the existing infrastructure.*
- Is your software cloud native? Which vendors (AWS, Azure, GCP) do you support?
- Do you support hosted/private cloud or on-premise deployment?
- What is the solution's ability to support real-time deployment?
- How does it support horizontal and vertical scaling?
- How does it offer high availability and how are failover procedures handled?
- How does it support backup and recovery procedures?
- How does it handle archiving?
- How do you support edge capabilities? Do you offer on-premises deployments?
- Is the solution validated with the standards of W3C and HTML5 to enable browser independence?
- Does it track the lineage of all data objects and code, showing upstream sources and downstream consumption?
- How does development occur e.g. introducing changes to core components, adding extensions?
- Is it possible to test reconfigurations, upgrades, and extensions before they are put into production?
- What are the software and hardware prerequisites?

# Project Execution, Services, and Support

Understanding how potential solution providers implement projects will allow you to anticipate time to value and create a high-level roadmap for implementation. The potential solution provider should provide the resources to ensure continued success. As successful implementations require both the right technology and the right support, this section helps you evaluate how well supported your team and organization will be when adopting an Industrial DataOps solution.

## Questions to evaluate a potential solution provider:

- Can you describe the go-live period between proposed solution validation/operational deployment, and final acceptance/beginning of any maintenance and support agreements?

*Expert Tip: This functionality saves time for data engineers and makes data discoverable for domain experts and other data consumers.*
- What maintenance and support do you offer during and after implementation?

*Expert Tip: The proposed solution provider should have a designated customer support representative to ensure project success.*
- What does a typical project implementation process look like? What support is available?
- What level of services do you typically provide?
- Please describe how your skilled experts will interact with our (the client's) in-house experts to maximize the benefit from collaboration?
- How do you enable/support search in the proposed solution? Can you provide documentation?
- How does the solution support documentation and how is it made accessible?
- What training programs are included and offered? What is typical?
- How do you ensure that competence is built within our (the client's) organization?

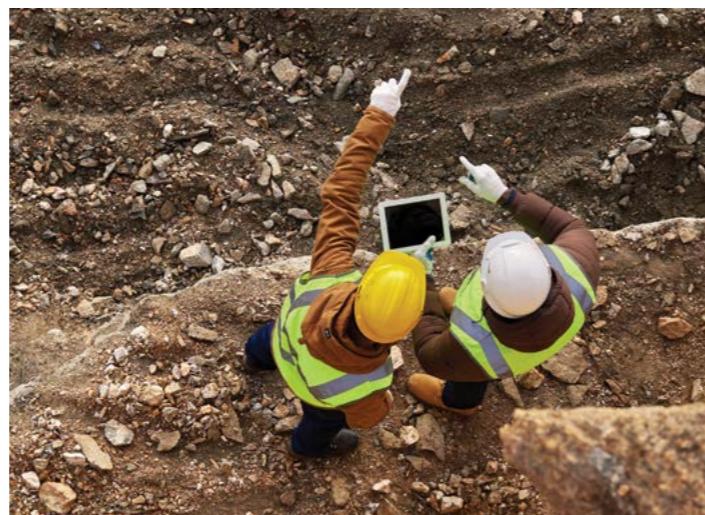
*Expert Tip: Building competence within your organization is an important step towards digital maturity. Your solution provider should be enabling these competencies. Otherwise your organization runs the risk of being in a service-based relationship with the solution provider.*
- What resources and support are provided during this period?
- What standard support do you provide in problem resolution? Do you offer varied support levels?

# Security

With the importance of security always increasing, the potential solution provider must be ready to meet the needs of your organization. This is not intended to be a comprehensive security list, as your IT department has likely developed its own security requirements for new software products. However, here are some of the key issues to consider.

## Questions to evaluate a potential solution provider:

- What is your company's strategy for penetration testing and third-party assessments?
- How does the solution maintain an audit trail of all data manipulation?
- How does it offer monitoring and statistics of backbone components?
- How do you ensure that we (the client) have access to our own data?
- How is high availability maintained for security, access, and governance?
- How do you support revocation of access at both user and group level?
- When and how data are encrypted in the proposed solution?
- What is the solution's capability with regard to access control? What is the granularity?
- Does it support groups for access control?
- Can authentication requirements be customized in the proposed solution?
- How does a user report suspicious activity related to data points?
- Can users be assigned special roles to fix or disapprove reported suspicious data points?
- Does the solution support ISO standards (or other standards as required)?
- How does the solution track the chain of custody?



# Usability

Successful rollout of any software solution depends on user engagement. Poor usability is a leading cause of poor product adoption. In order to make data discoverable and usable for all data consumers, the proposed solution must be intuitive and have well designed user interfaces that do not require strong coding backgrounds to operate. In addition, one of the largest complaints of data scientists is accessibility to data, even when centralized into a data lake. The potential solution provider needs to support both of these user groups to truly make data usable.

## Questions to evaluate a potential solution:

- Are users able to navigate through different parts of the proposed solution without help?
  - Expert Tip: Asking for a product demo is helpful when trying to assess this topic.*
- Do users see and feel the solution responding in real time?
- How many concurrent users does the solution support? And is the environment collaborative?
  - Expert Tip: As your Industrial DataOps solution gains users, your organization should be striving to increase adoption further, driving use case development across multiple departments.*
- How does the solution handle error messaging? Are errors easily interpreted by users?
- Can users easily refine search results?
- Can users create data pipelines without IT assistance and without deep training in data engineering, SQL, or production processes? Do you provide a graphical user interface for pipeline creation?
- Can users execute other tasks during the execution of jobs? Are users alerted when jobs are complete?
- How do you ensure fast search results are returned to users?
- How do users report errors, bugs, service failures, and requests for new services or extensions to existing services?

# Software Maintenance

This section is designed to give you an understanding of the upkeep required after a solution has been implemented. Reliability is another important factor in product adoption. Improvements and enhancements to the proposed solution should not result in unexpected downtime, nor should the solution require a high level of manual support to ensure proper operation.

## Questions to evaluate a potential solution provider:

- How often do you release improvements to your products? Do you have major and minor release cycles?  
*Expert Tip: As your organization requires, be sure to understand the different management requirements between on-premise, private cloud, and public cloud offerings.*
- Are clients entitled to all product upgrades with the base software? When are upgrades required?
- How are clients notified about both scheduled and unscheduled maintenance/downtime?
- How are new versions/updates managed?
- What level of availability and uptime do you guarantee e.g. 99.5%? How do you track system uptime?



# Future Development

Ensure that the potential solution provider's roadmap is aligned with your organization's goals. Seeing their top technology development priorities will provide you with clarity on the product direction and it can continue to support your organization's growth.

## Questions to evaluate a potential solution provider:

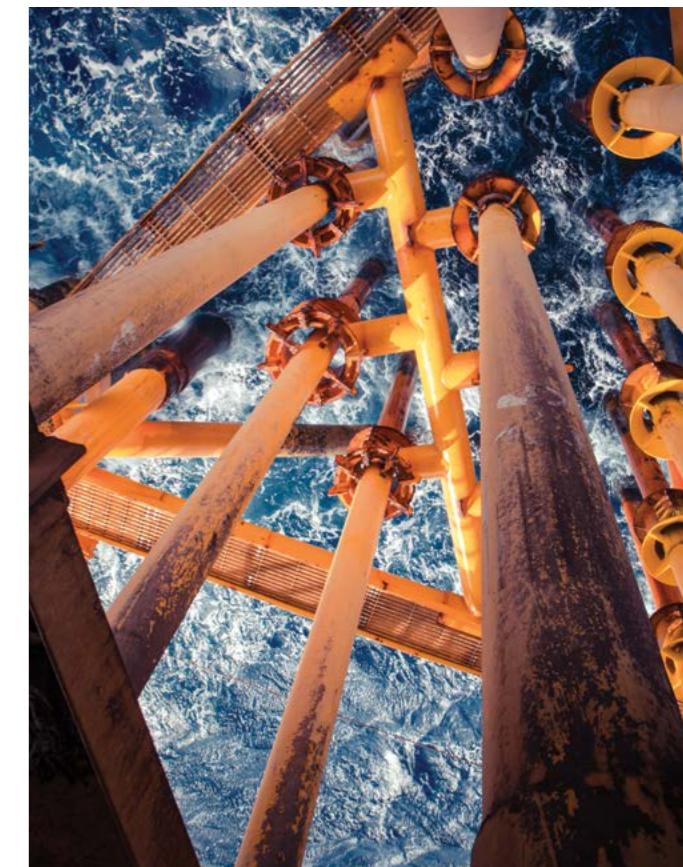
- Can you provide a short-term (6–12 months) and long-term (2–5 year) product roadmap?
- What is your approach to developing new products and the possibilities for developing customizations/extensions?

# Pricing Model

To date, price convergence has not yet taken place in the industrial software industry. Asking the high-level questions to understand the initial price (including services) required to get started will be valuable when assessing potential solution providers. In addition, Industrial DataOps solutions are designed to scale, so it's also important to understand the levers of pricing when data sources, users, and use cases start to increase.

## Questions to evaluate a potential solution provider:

- How do you price the product? How does your pricing model support increasing use case and product adoption?
- What factors do you predict will be the main cost drivers for your product and services?



As mentioned above, the purpose of this document is to provide a guideline to building an RFP for your Industrial DataOps solution, so you can achieve both current and future project success.

Current technology and functionality is often given the utmost importance, but having a solution that is easily adopted across your organization is equally important.

Industrial DataOps software, like all industrial software, can turn into shelf-ware if the solution is unfriendly and users are unable to effectively apply the available technology. This will inevitably lead to underperformance against your expected ROI.

Lastly, Industrial DataOps solutions are designed to become an integral part of your daily operations, so it's crucial to work with an experienced solution provider whose product roadmap shows they will grow with your future needs.

**With all of this in mind, you now have the knowledge to build an RFP for an Industrial DataOps solution that will enable your organization to extract significantly greater value from your data.**

# About the Authors



**Petteri Vainikka**

**Vice President, Product Marketing, Cognite**

Petteri's 20-year professional career spans consumer and enterprise SaaS technologies, where he has always found himself at the intersection of emerging transformational technology development and its commercial applications for customer value. Prior to Cognite, Petteri worked in senior product management, marketing, sales, and general management positions for companies such as Sumea, Rovio, Cxense, and Ardoq. Petteri has a master's degree in technology from Aalto University in Helsinki.



**Alexandra Teachout**

**Director of Product Marketing, Oil & Gas, Cognite**

Alexandra has spent over 15 years as a product marketing professional for Enterprise SaaS companies. With a background in location-based data, Alexandra has focused on operationalizing data for maximum value across multiple industries including transport and logistics, automotive, and smart cities. At Cognite, she leads the go-to-market activities for Cognite's Oil & Gas business with a specific focus on powering the energy transition.



**Jonathan Higgins**

**Director of Corporate Communications, Cognite**

A copywriter at heart, Jonathan's background spans government communications in the United States, academic and policy writing in Germany, and private sector PR and reputation management in Asia and the Nordics. Over the last 15 years, he has driven multi-market communications programs for global telecommunications, technology, automotive, and food & beverage giants; and has written regularly on behalf of many of their top leaders. At Cognite, he leads corporate communications. He is also editor of the industrial transformation news platform, Ignite.



## Gabe Prado

### **Director of Product Marketing, Power & Utilities, Cognite**

With a background in data acquisition, test and measurement, DataOps, and AI & ML systems for heavy industry, Gabe has spent his career observing and helping to solve the challenges that major industries face during their digital transformation journeys. At Cognite, he strategizes and manages key go-to-market activities for Cognite's Power & Utility business and has led research and benchmarking efforts to relate industrial data operations with digital maturity.



## Kiran Dhillon

### **Senior Director of Product Marketing, Cognite**

Kiran's career spans digital strategy across consumer technologies to enterprise SaaS. At Cognite, Kiran leads product definition and strategic vision for Cognite's suite of products and solution sets. Prior to Cognite, Kiran led product strategy and digital marketing at Ogilvy & Mather, Young & Rubicam, Opera Software, and Cxense. Kiran has held various positions in international markets such as India, Norway, and the US.



## Petter Jacob Jacobsen

### **Vice President, Corporate Development, Cognite**

Petter Jacob started his career as a field engineer in oil and gas before moving into data science, a combination that came in very useful when he joined Cognite, almost from its infancy. At Cognite, Petter Jacob has been involved in various digital transformations in large industrials and was leading the go-to-market for Oil & Gas before moving into his current role in Corporate Development. Petter Jacob draws on his experience from the industrial and digitalization frontline to contribute to this book.



## Ben Skal

### **Director of Product Marketing, Manufacturing, Cognite**

Ben has spent the past 10 years working in the manufacturing industry. He started his career gaining hands-on experience as a process engineer in the steel industry, followed by project managing automation solutions for process industries. He experienced many of the challenges manufacturers have to operationalize their data, and is focused on helping these organizations adopt new technologies into their workflows and transform their operations.



