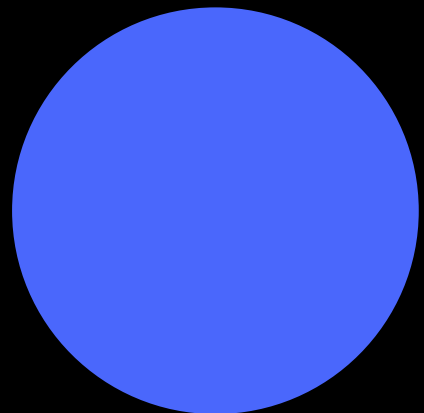
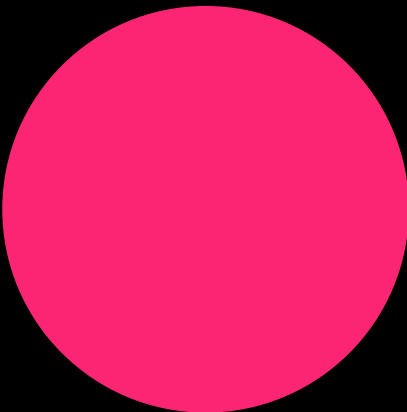
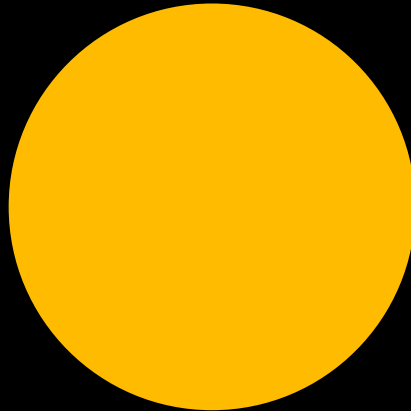
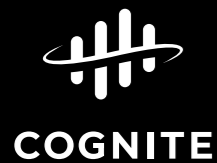




The Definitive Guide to Generative AI for Industry



How Digital Mavericks are Redefining
the Rules of Digital Transformation



The Definitive Guide to Generative AI for Industry

How Digital Mavericks are Redefining
the Rules of Digital Transformation

©Copyright, Cognite, 2023 - www.cognite.ai →

“The seismic new AI frontier offers an unprecedented opportunity to harness technology to transform industry at a scale that impacts national GDPs - but, to be successful, AI needs context. AI needs a strong data foundation. Now is the time to place bold bets in digitalization and usher in tremendous simplicity, scale, and a clear path to autonomous operations. The Definitive Guide to Generative AI for Industry is a step-by-step guide for taking advantage of this enormous opportunity within your organization.”

MOE TANABIAN
CHIEF PRODUCT OFFICER
COGNITE

“Safe, secure, hallucination-free generative AI is critical to paving the road to sustainable and profitable global energy supply and manufacturing excellence. This is the moment Cognite was built for. A culmination of our extensive Industrial AI experience, The Definitive Guide to Generative AI for Industry provides a comprehensive how-to for transformation leaders looking to drive actual bottom-line impact.”

GIRISH RISHI
CEO AND CHAIRMAN
COGNITE

“Generative AI is fundamentally reshaping operational processes and it will be the digital mavericks who will be early adopters of this technology. Celanese has set the standard in industrial innovation with its Digital Plant of the Future; and by working with Cognite to accelerate AI implementation into our digital transformation strategy, it will allow us to raise the stakes once again and solidify Celanese as a leader in advanced manufacturing.”

BRENDA STOUT
VICE PRESIDENT MANUFACTURING
CELANESE

“Cognite’s work so far with LLMs, grounded by its Industrial DataOps platform, is remarkable – with even greater value on the horizon in the form of not just data management, but information orchestrated by real, thoughtful AI”

JOE LAMMING
INDUSTRY ANALYST
VERDANTIX

Contents

→ 06 Foreword Digital Mavericks: The Individuals Leading the Industrial Transformation	→ 38 Chapter 3 Defining Industrial Knowledge Graphs	→ 78 Chapter 7 Contextualization, Data Models, and Digital Twins	→ 122 Chapter 11 Impact of Gen AI on Asset Performance Management
→ 18 Executive Summary The Four Things You Need to Know about Generative AI for Industry	→ 44 Chapter 4 Making Generative AI Work for Industry	→ 92 Chapter 8 AI Is the Driving Force for Industrial Transformation	→ 136 Chapter 12 The Next Phase of Digital Industrial Platforms
→ 24 Chapter 1 Short History of AI in Industry	→ 58 Chapter 5 The Industrial Data and AI Problem	→ 104 Chapter 9 Use Cases for Industrial Generative AI	→ 154 Chapter 13 Tools for the Digital Maverick
→ 34 Chapter 2 Demystifying LLMs	→ 68 Chapter 6 A Robust Contextualization Engine	→ 116 Chapter 10 Impact of Gen AI on Robotics and Autonomous Industry	

Digital Mavericks:

The Individuals Leading
the Industrial Transformation



“No one ever got fired for choosing IBM.”

Once considered a safe bet for capital projects, purchasing services from IBM offered large enterprises a certain guarantee of results and peace of mind for complex, multi-year initiatives. This thinking was acceptable when business reinvention was measured in years and decades. However, in today's new era of hyper-digitalization and generative artificial intelligence (AI), maintaining this long-standing business cliché carries increasing, compounding risk. Buying 'safe' often comes at the cost of speedy innovation and long-term differentiation, where asset-heavy industry has much to learn and where a new class of business leaders recalibrates the risk vs. reward of these strategic technology investments.

As Gartner puts it in their Maverick predictions series: “Best practices don't last forever—you need to look toward next practices, too. Challenge your thinking by considering less-obvious developments.”¹ If leading industrial companies prioritize their long-standing business playbooks in this new era of AI, they risk missing their 'iPhone' moment, where making certain bets in digitization (at the right time) pay off in an outsized way, ushering in tremendous simplicity, scale, and a clear, valuable path to autonomous operations.

There is no more apparent opportunity to implement this strategy than by applying emerging innovations in generative AI technology such as ChatGPT and GPT-4. Generative AI possesses the potential to fundamentally reshape how knowledge workers and subject matter experts interact with data and operational processes. In fact, Gartner estimates that “by 2030, 75% of operational decisions will be made within an AI-enabled application or process.”² Connecting the dots between the business and the technology, creating the opportunity, and driving the required change management takes a clear vision and a strong immunity and resilience to failure.

Industrial organizations can no longer afford not to take risks. But which heroes are central to business transformation and are staking their careers (and driving meaningful change) based on the truth that digitalization across industry remains deficient and that a smarter, more autonomous industrial future requires a new playbook?

The Rise of the 'Digital Maverick'

Characterized by strategic, big-picture thinking, relentless determination, strong business understanding, technical prowess, and sometimes a chip on their shoulder, the digital maverick is a new breed of industrial leader who is critical for taking an industrial digital transformation program and making it truly meaningful and valuable with the advent of industrial AI.

While other members of digital operations teams think about deploying operational use cases and analytics, the digital maverick is thinking about building new long-term business capabilities. Instead of searching for keywords such as: “what is an industrial chatbot?,” “good data management platforms,” or “digital operations dashboards,” the digital maverick craves and drives clarity on new ideas such as “ChatGPT for operations,” “retrieval augmented generation,” “AI and contextualized data,” and “autonomous digital factory.”

They also know from experience what has not worked well enough to deliver meaningful value from previous attempts at industrial AI—either from seeing first-hand more noise than value (unreliable models, too many false positives, constant retraining), or outright failure of AI deployments to create meaningful change in operator workflows. The digital maverick knows there's a solution to the age-old industrial challenge of operating complex environments with limited insights and sub-optimal decision-making.



1. www.gartner.com/en/podcasts/thinkcast/behind-the-research-maverick-research →

2. Gartner Report: Maverick Research: Data and Analytics Roles Will No Longer Be a Priority →

Q&A with a Digital Maverick:



Ibrahim Al-Syed, Director
of Digital Manufacturing

To truly understand this digital maverick mindset and perspective, we spoke with one of these innovative thinkers, Ibrahim Al-Syed, the Director of Digital Manufacturing at Celanese Corporation, a global chemical and specialty materials company.



Q: You've been working in the area of digital transformation for most of your career; can you tell us a bit about your journey.

Ibrahim: For me, it was a very natural transition. Earlier in my career, I worked at an integrated refinery and chemical plant. I was involved in a program to improve operational processes. This meant we were looking at design issues, hardware issues, and the operating model to improve our asset integrity program. What became very apparent while working in that role was that we needed to have easier access to more data in order to quickly get insights for smarter troubleshooting and reliability solutions.

I still remember the conversation I was having in my boss's office where we drew this concept of a 'digital twin' on a whiteboard. I said, "What if we can have assets at the center and, in just one click, know everything about that asset, what has happened in the past, what is happening now, and what could happen in the future?"

Then we started looking at what technologies we would use to reach that goal. That was the first time we started looking at a concept that would leverage data as a step change in performance for integrity and reliability.

Q: What key challenges have you experienced when adopting new operational technology?

Ibrahim: While working both in Singapore and the U.S., I've seen the same pain repeatedly. There's so much siloed technology out there, with individual solutions for almost every operational problem ... AI, digital twins, too many UIs. Something needs to be fixed if it is still easier to pass paper around than trying to go from tool to tool to find data and prepare it to fit in another tool or work process. We have to move away from situations where a single event at a plant generates activity that has limited actual value on the asset.

Here's where my idea of human-centered digital transformation comes into play. How can we better architect the interactions between solutions and technology with the data and processes we need to empower our people? It's how you transform their ways of working so that they are optimized, inspired, and work at their maximum potential. So you have to design backwards from the people, starting with a clear vision of where you want to go and what it's worth.

Q: When you look at your progress so far, what approach are you taking to help you get to this human-centered future?

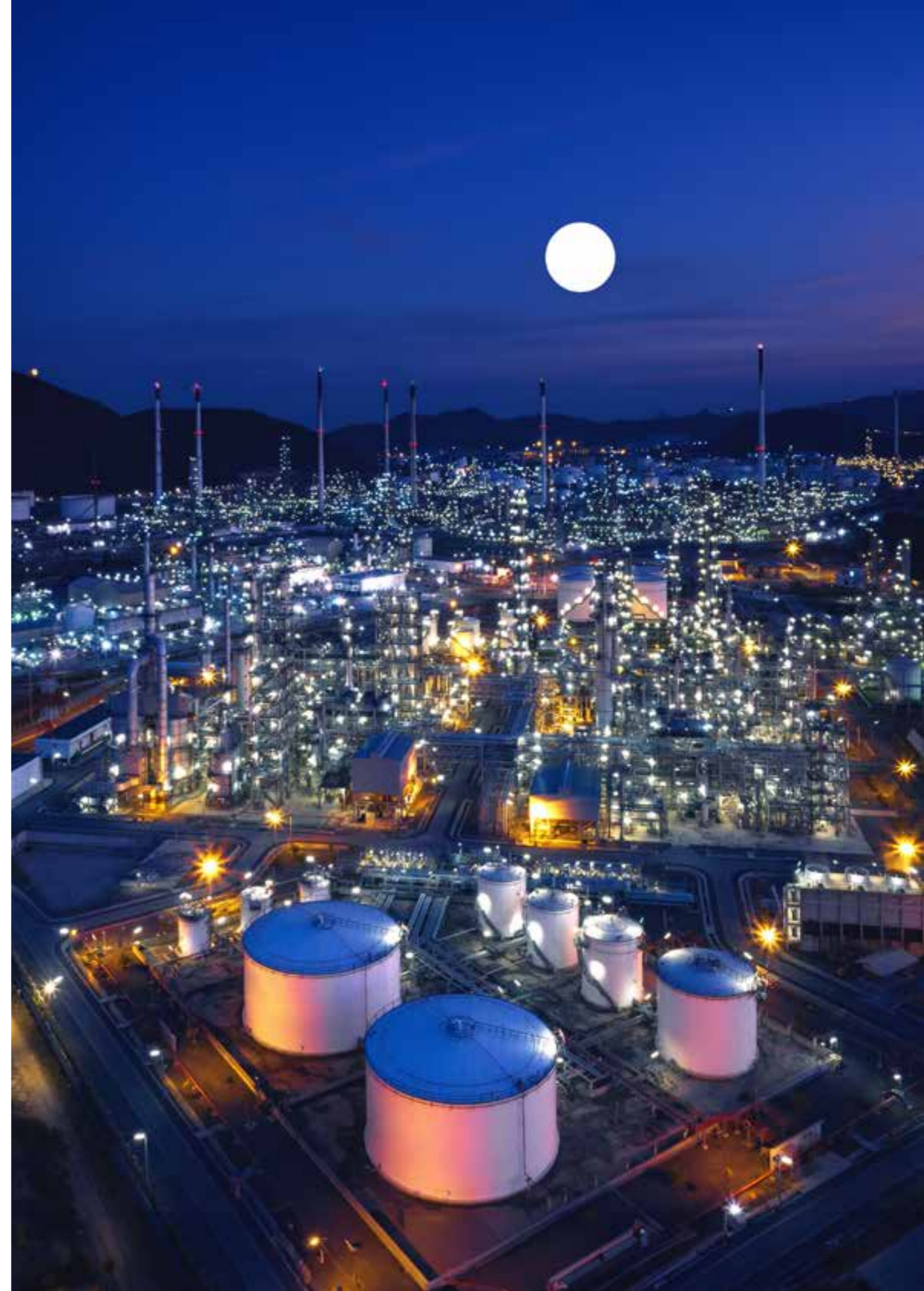
Ibrahim: We want to reach a point at which we operate facilities by exception. But that level of maturity requires a new level of human-centered use cases and an advanced way of using and propagating data.

So when you build any new product, you do market research and understand customer feedback to create something people want to use. You have to do the same—what I call user research—in digital transformation to give these new business capabilities the best chance to get adopted at the plant level.

When our people are involved in the design process, their experiences allow us to model the data to propagate it seamlessly. What I learned about change management is that people will accept some of the 'broken stuff' as long as they own, lead, and drive the journey.

Q: Looking ahead, where do you see generative AI for industry fitting with the need for more intuitive operator experiences around data and analytics? What use case/business capability is most compelling?

Ibrahim: Generative AI is a game changer for our industry. We are finally to the point where we can get 'steel and assets to talk to us,' which can help companies run and operate facilities by exception. You can have insights and generated actions before you even have time to click a button. It is important to note that highly experienced workers are retiring, and their experience and knowledge can often be lost. These are the workers who can hear when a machine bearing needs grease, or feel when a machine is vibrating excessively and not running correctly. With talent shortages and retiring employees, generative AI promises to maximize overall equipment effectiveness in production environments.



Digital Mavericks: This is Your Moment



Everything these digital mavericks have been implementing with incremental gains to date (Industrial DataOps infrastructure, skillset investments, and other foundational capabilities) have become **10X more valuable** almost overnight as industry adopts high-trust, secure, hallucination-free generative AI to harvest more business value from digital operations.

The inflection point could not be more real, obvious, or urgent from both a technology and strategic perspective. True digital mavericks know that the game has been changed forever and that they carry a new responsibility to drive and govern a new playbook that includes a new set of guidance:



1. Realign digital KPIs and skill sets to key business drivers

Digital mavericks know that their organization's charter and KPIs must be ever more tied to business value and operational gains. Instead of deploying proofs of concepts, their KPIs must reflect business impact, successful scaling, and other product-like metrics:

- How can business value be measured, quantified, and directly attributed to the digital initiative? What 'portfolio' or 'menu' of value is being developed?
- How is adoption measured and user feedback implemented in the feedback loop? How many daily active operations users are actually in the tools delivered? How is the workflow changing as a result?
- How much do these business applications and solutions cost to deploy and maintain? What do they cost to scale to the following asset, site, or plant?
- What new business capabilities for multiple stakeholders are gained due to deploying a solution? Are these capabilities short term, or will they persist (and drive value) over the long term?

Additionally, the portfolio of skills needed to develop and run analytics departments is also positioned to change. Gartner estimates that "by 2030, the number of traditional descriptive analytics dashboards will decrease by more than 50% in most modern digital businesses."³ Generative AI will have successfully abstracted away the traditional complexity of creating and managing specialized analytics in favor of more business-ready insights.

The clear message to legacy CIOs and want-to-be mavericks? Stop investing in IT and data and analytics skills and focus more on connecting AI-powered use cases to business impact.

³ Gartner Report: Maverick Research: Data and Analytics Roles Will No Longer Be a Priority →

2. Land and expand with business value faster than the competition

Technology can change in an instant, but operational KPIs are evergreen. Tackling digital transformation holistically across the enterprise has yet to prove to move the needle fast enough to deliver a meaningful competitive edge in dynamic markets. Instead, digital mavericks are pursuing 'land and expand' approaches that integrate new technology in the context of high-value use cases, with clear pathways from initial test concepts into scaled (and 10x more valuable) deployments.

Whether the team is focused on driving down OPEX, increasing production capacity from shorter turnarounds (TARs), or eliminating waste from a production process, think beyond siloed processes. Can this new use case be deployed across sites? How valuable is it compared to other business requests? What is the cost and effort to scale?

Here, the digital maverick puts a premium on flexible technology, approaches, and teams so that they can be prepared to shift strategy—either to take advantage of market opportunities or new technology such as generative AI—at a moment's notice. For example, GPT appeared in a decent state of maturity within just a few quarterly business cycles, disrupting product roadmaps and, just like the pre-internet days, sending the world to a new level that will never come back. Organizations equipped to mobilize and execute on these trends can capitalize on market opportunities much faster than their competitors.

3. Challenge traditional thinking around DIY (do-it-yourself) projects and technology

Digital mavericks have always been tempted to go down the DIY path. Still, more have started to realize the significant opportunity costs that come with this approach's strategy, implementation, and change management—especially when it comes to time. Is your largest competitor looking for off-the-shelf generative AI components for their digital transformation while you invest resources in more cloud-based building blocks?

Instead of making a name based on 'completeness and sophistication of tech stack,' consider building a reputation based on 'time to scaled value,' a far more critical metric. Here are a few other long-standing myths that digital mavericks are starting to challenge:

- They recognize that DIY is not usually the least expensive option when considering headcount and long-term total cost of ownership of infrastructure and deployed solutions.
- They are wary of going all-in on one cloud provider's capabilities, knowing that the only way to avoid vendor lock-in is by investing in multi-cloud ecosystems that include DIY and SaaS.
- They understand that differentiation in technical and go-to-market abilities is a function of speed and agility, not having a potentially cumbersome, custom home-grown data platform.

- They know that SaaS is also a valid enterprise-grade path and appreciate that mass-market software spreads development risk and comes with self-service documentation, support SLAs, and other entitlements that minimize costs.
- They have seen that job security and department prestige is a product of delivering actual value in operational workflows, not being seen as a high-headcount Proof of Concept (PoC) factory with little quantifiable value to show.
- They acknowledge that their industrial company will never be in the software business and that, to be true innovators, they can't repeat pre-cloud, pre-SaaS-era bespoke software development practices.



The Four Things You Need to Know about Generative AI for Industry

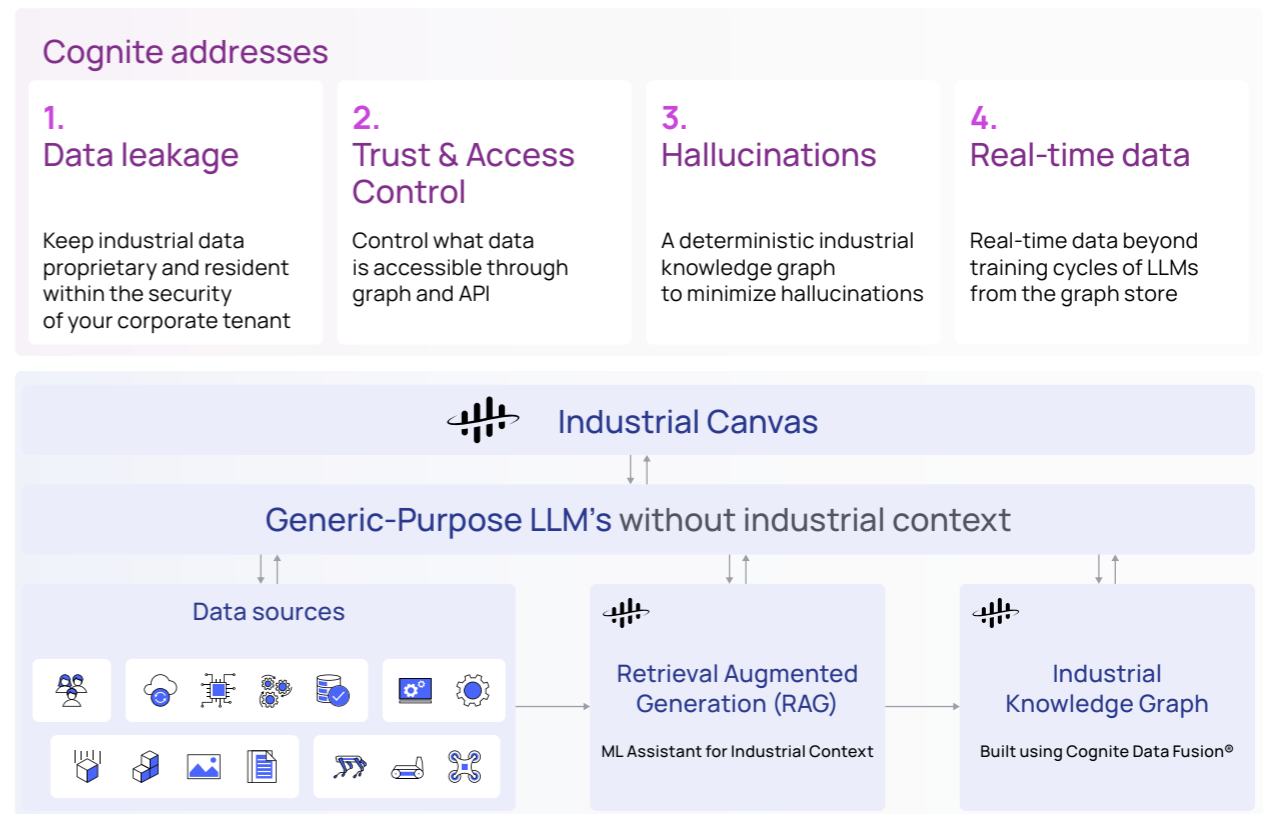
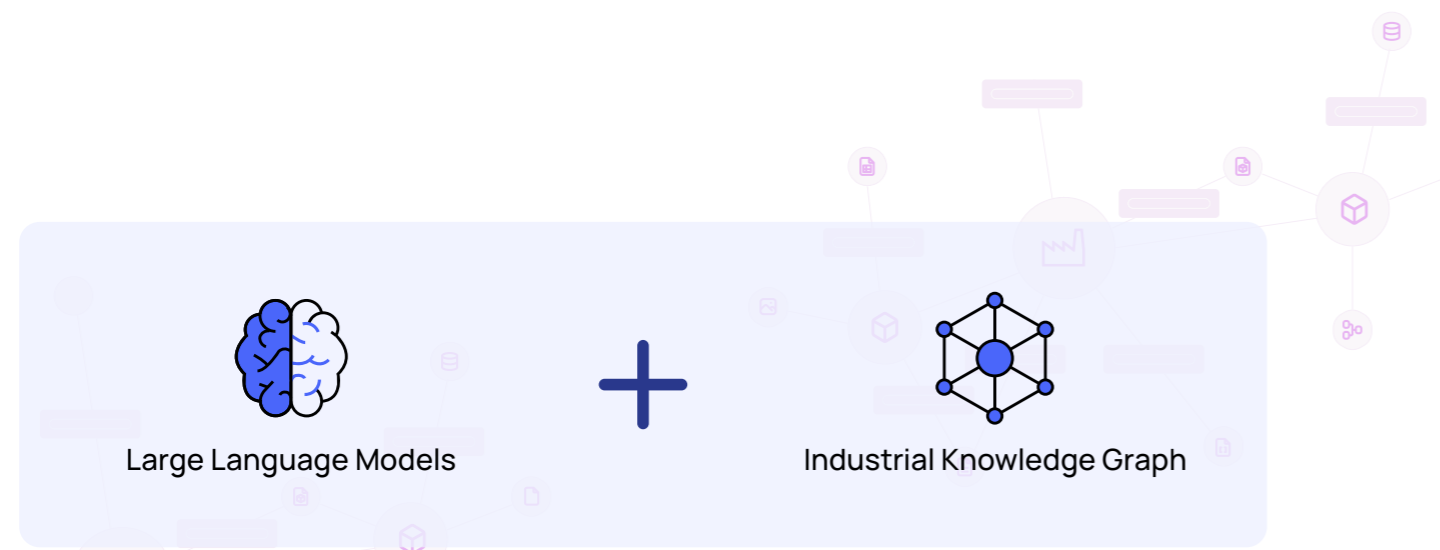


You're Receiving AI Messages from Everyone.

There is a lot of information out there (and, yes, we realize we've added a 184-page book to the onslaught). It can be a lot to take in. To help, we've boiled down generative AI for Industry to four key points. We will dive into more details on each of these points in the following chapters, but if you only read one part of this book, let it be this:

1. LLMs + knowledge graph = Trusted, explainable generative AI for industry

This is the simple formula to apply generative AI for industry. Your asset performance management is made intelligent and efficient by combining large language models (LLMs) with a deterministic industrial knowledge graph containing your operations data.



Cognite addresses

1. Data leakage

Keep industrial data proprietary and resident within the security of your corporate tenant

2. Trust & Access Control

Control what data is accessible through graph and API

3. Hallucinations

A deterministic industrial knowledge graph to minimize hallucinations

4. Real-time data

Real-time data beyond training cycles of LLMs from the graph store

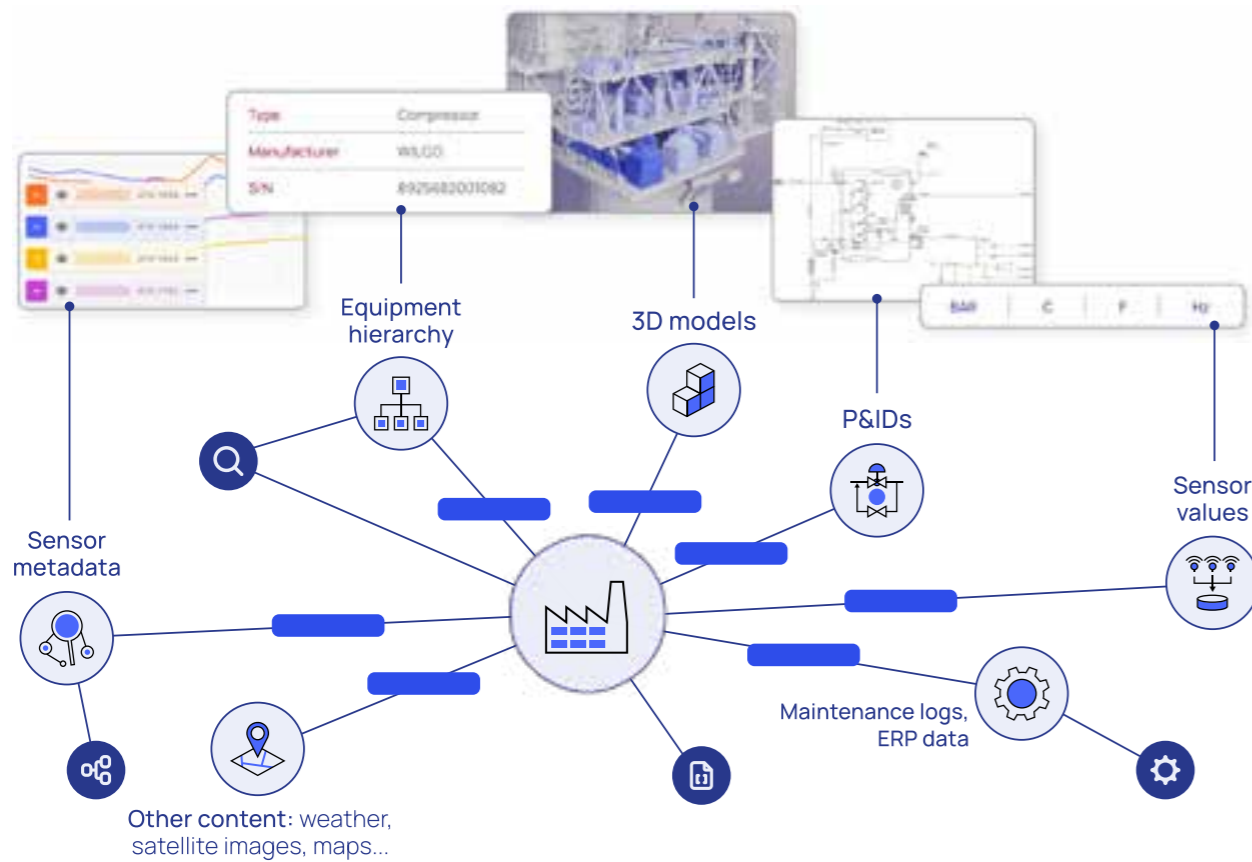
2. Generative AI for industry needs to be safe, secure, and hallucination-free

And with the previous formula, it is. You need a complete, trustworthy digital representation of your industrial reality (industrial knowledge graph) for LLMs to understand your operations, and provide deterministic responses to even the most complex questions.



3. To apply generative AI in industrial environments, the ability to prompt LLMs with your operational context is everything

This means having a deterministic industrial knowledge graph of your operations, including real-time data. You need a solution that delivers contextualized data-as-a-service with data contextualization pipelines designed for fast, continuous knowledge graph population.



Contextualize events to assets

E.g. connect shop orders, work orders, and alarms, to correct site, prod. line and equipment



Contextualized documentation

E.g. connect tags in P&IDs to associated assets, time series, work orders, and documents etc.



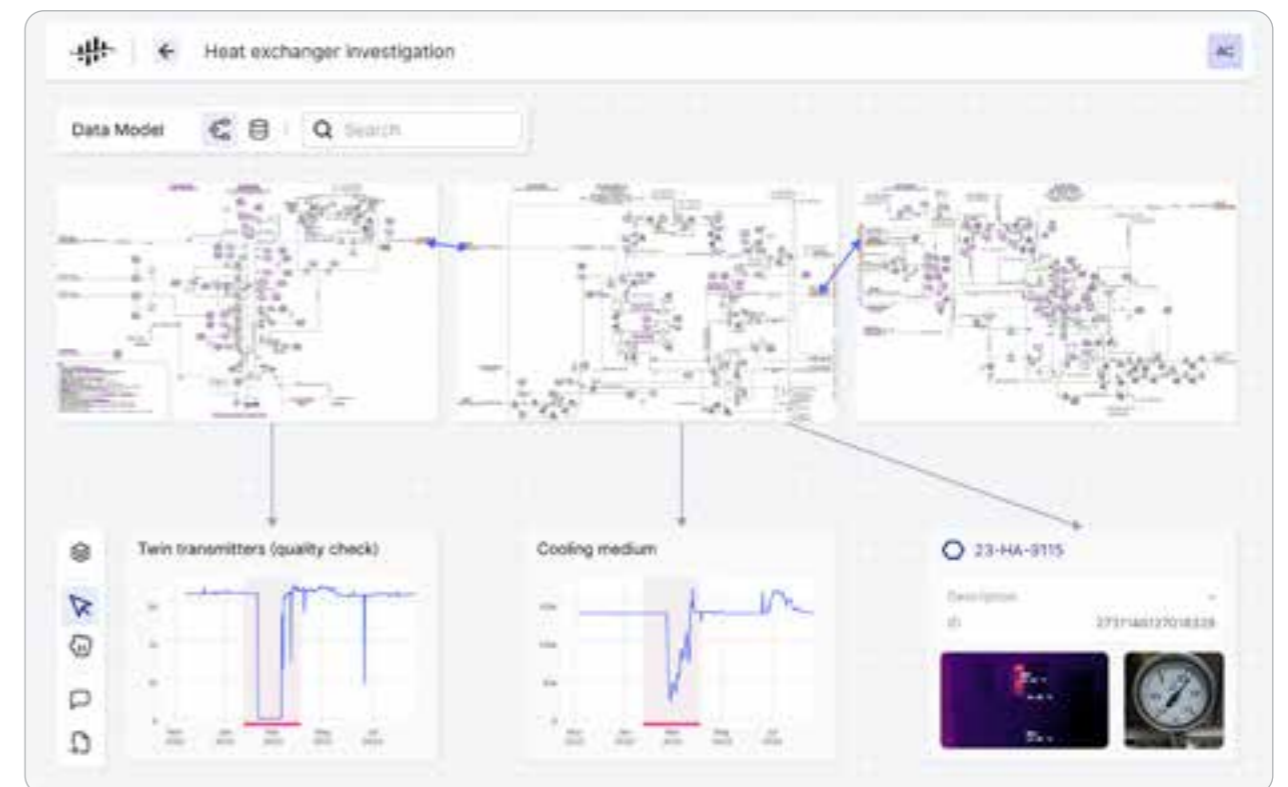
Contextualized 3D models

E.g. filter for and visualize where work orders are located on plant



4. While generative AI itself is undeniably transformative, its business value is in its application to the real-world needs of field engineers

Generative AI can already be applied today across field productivity, maintenance planning, and robotic automation, but only with a platform that delivers essential AI features that enable simple access to complex industrial data for engineers, subject matter experts, data scientists, and more.



Short History of AI in Industry



Introduction

As we delve into the state of industrial AI, we must acknowledge that AI's impact radius has grown substantially, expanding its influence beyond individual machines to encompass entire production systems, reshaping the way industries operate.

Artificial Intelligence	Machine Learning	Deep Learning	Generative AI
<p>1956</p> <p>Artificial Intelligence is the field of computer science that seeks to create intelligent machines that can replicate or exceed human intelligence</p>	<p>1997</p> <p>Machine Learning is a subset of AI that enables machines to learn from existing data and improve upon that data to make decisions or predictions</p>	<p>2012</p> <p>Deep Learning is a machine learning technique in which layers of neural networks are used to process data and make decisions</p>	<p>2021</p> <p>Generative AI can create new written, visual, and auditory content given prompts or existing data</p>
<p>Examples:</p>			
	<p>Educate a statistical model on the characteristics of a failure within the data, and provide it with relevant supporting data to enable it to identify patterns leading up to such failures.</p>	<p>Enable the statistical model to not only detect failures but also autonomously determine which data is significant in accurately predicting those failures.</p>	<p>Empower the operator to make natural language inquiries on previously unassigned data for analysis, such as, "Which pumps are at the highest risk of exhibiting signs of failure in the next week if production capacity is raised by 10%?"</p>

For most of industrial history, technological progress primarily affected the physical aspects of business (better machines, more efficient factories). Now, we are witnessing remarkable advancements in machine learning techniques and computing power. These advancements have enabled AI to transcend its previous limitations, propelling it into organizational decision-making in a way not thought possible even a year ago. Today, AI is not merely an add-on or a supplement to industrial processes; it is quickly becoming an integral part of modern industrial operations.

From sophisticated robotics to intelligent sensors, machines are now equipped with AI-powered capabilities that enable them to adapt, learn, and optimize their performance. AI algorithms can now analyze vast amounts of data collected from various sources to optimize operations holistically, identifying patterns, predicting failures, and making intelligent decisions.

AI is optimizing production schedules, minimizing waste, streamlining resource allocation, and providing a better understanding of market dynamics and customer behavior. And AI-enabled predictive modeling and simulation techniques enable organizations to mitigate risks and maximize returns on investment.



AI-driven insights are quickly becoming the foundation for business decisions that continually improve operational efficiencies, drive innovation, enhance customer experiences, and identify new business opportunities.

Integrating AI into machines, production systems, and organizations has brought about a paradigm shift in how industry operates. AI is soon to be the backbone of the modern industrial landscape. In the following sections, we will explore the latest advancements, trends, and case studies that highlight the profound implications of AI in reshaping industry and propelling it toward a smarter, more sustainable future.

Predictive Maintenance: Two Decades in the Making and the Role of AI

Predictive maintenance, the practice of utilizing data-driven analytics to predict and prevent equipment failures before they occur, has been a goal pursued by industry since the early 2000s. Its potential to optimize maintenance operations, reduce downtime, and improve overall equipment effectiveness has long been recognized.

Initial machine learning methods, which relied on statistical analysis and rule-based approaches to detect anomalies and trigger maintenance actions, were often limited by their reliance on pre-defined thresholds and a lack of flexibility to adapt to dynamic operational conditions. The



introduction of Deep Learning (vintage 2012) is what helped to overcome these initial limitations because the transformer model that powers Deep Learning neural networks made it possible to use AI to identify and determine which data is most valuable in predicting future outcomes. In contrast, early machine learning relied upon the data scientist to identify and determine the best data features with predictive value. This ability to automatically learn from data and adapt to changing conditions has significantly enhanced the accuracy and effectiveness of predictive maintenance systems.

However, as discussed in detail in a review published in the CIRP Journal of Manufacturing Science and Technology⁴, a lingering challenge of predictive maintenance is the availability and accessibility of data. Traditional maintenance practices often relied on scheduled maintenance routines or reactive approaches, where maintenance was performed after a failure occurred. As a result, there is a lack of historical data on equipment performance and failure patterns.

To solve the challenge of data availability, we saw the rise of industrial IoT and sensor technologies that enabled real-time data collection from machines, making it possible to monitor their condition and performance continuously. Unfortunately, this feeds into the accessibility of data problem: too much information.

Modern AI, when paired with a consolidated and contextualized data foundation, can overcome many of the challenges predictive maintenance has faced in the past. AI-powered algorithms can now analyze complex data sets, including sensor readings, environmental factors, operational parameters, and even unstructured data such as maintenance logs or operator feedback. AI systems can identify early warning signs of equipment degradation or failure by employing advanced techniques such as anomaly detection, pattern recognition, and predictive modeling.

The use of AI in predictive maintenance has also paved the way for new approaches, such as prescriptive maintenance. By combining predictive analytics with optimization techniques, prescriptive maintenance systems can recommend the most efficient maintenance actions, considering factors such as equipment criticality, resource availability, and cost considerations. AI enables organizations to predict and prevent failures and optimize maintenance strategies for improved overall operational efficiency.

Integrating AI and machine learning has played a pivotal role in unlocking the true potential of predictive maintenance, enabling industries to transition from reactive practices to proactive, data-driven approaches. As AI continues to advance and industries embrace its potential, the future of predictive maintenance holds even more promise, with the ability to transform maintenance operations, enhance asset performance, and drive the growth of intelligent, resilient industrial ecosystems.



4. www.sciencedirect.com/science/article/pii/S1755581722001742 →

Hybrid AI: The Convergence of Human Expertise and Machine Intelligence

The significance of hybrid AI in industry lies in its ability to bridge the gap between data-driven insights and human understanding. It acknowledges that while machines excel at processing large volumes of data, identifying patterns, and performing repetitive tasks with precision, they still need human experts' contextual understanding, intuition, and creativity. Combining AI algorithms' analytical power with human domain knowledge and decision-making skills, hybrid AI aims to create a symbiotic relationship that harnesses the strengths of both humans and machines.

One key aspect of hybrid AI is its role in addressing the 'explainability' of AI systems. Traditional black-box deep learning models often lack transparency, making it challenging for humans to understand and trust their outputs. This more informed reality of AI in industry is driving the future of hybrid machine learning, a blend of physics and AI analytics that combines the 'glass box' interpretability and robust mathematical foundation of physics-based modeling with AI's scalability and pattern recognition capabilities. By combining physics and AI analytics, hybrid AI systems provide a more comprehensive and understandable rationale behind their recommendations, increasing trust and facilitating human acceptance of AI-driven decisions.

Moreover, hybrid AI enables continuous learning and improvement. Human feedback and interventions can be integrated into AI models, enhancing accuracy, adaptability, and generalizability. Human experts can refine AI algorithms by providing domain-specific knowledge, validating results, and correcting biases or inaccuracies. This iterative collaboration between humans and machines strengthens the AI models over time, resulting in more robust and effective solutions.

Hybrid AI is best suited for complex industrial process problems where a mathematical theory framework exists that can be used to teach a machine learning model that is then used on real-time data for predictions. The result is a high-confidence, tailored hybrid model combining strong domain knowledge (physics) with machine learning for cost efficiency and scalability.

By empowering humans with AI-driven insights and automation, hybrid AI enables industry professionals to make more informed decisions, identify hidden patterns or anomalies, and optimize complex processes. All of which will enhance decision-making and operational efficiency.





Generative AI: The Accelerator of Everything

Generative AI is a class of AI techniques that can generate new data, content, or solutions based on patterns and insights derived from existing data. Unlike traditional AI models that rely on pre-defined rules or training on large labeled data sets, generative AI models can learn from context-enriched data without explicit guidance, enabling them to create novel outputs that mimic the characteristics of the training data.

Generative AI models can ingest diverse data sets, including historical maintenance records, sensor data, work orders, and even unstructured data such as maintenance reports or equipment manuals. If this data can be augmented to include semantic, meaningful relationships (i.e. context), generative AI can recognize patterns and correlations to create detailed and optimized turnaround plans, for example, considering factors such as resource availability, dependencies between tasks, budget constraints, and safety considerations.

However, simply ingesting the diverse data sets without context will prevent generative AI from being deterministic (more of which will be covered in coming chapters).

Generative AI has emerged as a powerful catalyst, propelling industry towards new frontiers of efficiency, innovation, and operational excellence. Key to the wildfire that is generative AI is the accessibility of data it provides through human language interfaces and AI copilots.

With the advent of natural language processing (NLP) techniques and conversational AI, human operators can interact with generative AI models using everyday language, making the technology more accessible to a broader range of users.

Through copilots, operators can easily communicate their requirements, constraints, and preferences to the generative AI models. For example, an operator can specify the desired turnaround duration, prioritize specific maintenance tasks, or account for particular safety protocols. Generative AI models can then process this information and generate turnaround plans that align with the operator's objectives while optimizing various factors for efficiency and cost-effectiveness.

Copilots seamlessly integrate with human operators' workflows, providing real-time suggestions, explanations, and feedback during any industrial process. The operator fine-tunes the outputs based on domain expertise and the AI empowers operators to make faster, more informed decisions. This collaborative interaction between human operators and generative AI models enhances the quality and accuracy of the AI's results.

By automating and optimizing complex processes traditionally reliant on manual efforts, generative AI accelerates all industrial digitalization efforts we have embarked on up to this point and spurs new frontiers for digitalization we have yet to imagine. Generative AI stands to be the driving force behind industrial digitalization, revolutionizing operations and processes across various sectors.

But how do we make generative AI work for highly risk-averse industries?

The next chapter dives into how to enable a hallucination-free AI experience with your enterprise-wide data.

Demystifying LLMs



Understanding Large Language Models and their Application in Operations

Large Language Models, also known as LLMs, are perhaps the biggest buzzword since blockchain.

Unlike its complex and challenging-to-implement predecessor, business professionals can leverage LLMs without extensive prerequisites.

Whereas blockchain often requires specialized knowledge, technical expertise, and substantial investments in infrastructure, LLMs provide a much more accessible and straightforward solution. They are easily integrated into existing workflows and systems, enabling seamless adoption and integration with business operations.

At its core, an LLM is an AI model that utilizes deep learning techniques to understand, interpret, and generate human language, enabling businesses to leverage their capabilities for a wide range of applications.



LLMs are trained on vast amounts of textual data from diverse sources, such as books, articles, and the internet, allowing them to develop an understanding of grammar, context, and semantic relationships. These models can then generate coherent and contextually appropriate text based on user prompts or queries.

The history of LLMs can be traced back to the development of neural networks and deep learning algorithms. Early models, such as recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, paved the way for language understanding and text generation tasks. However, it was the advent of transformer models, with the introduction of architectures such as the GPT (Generative Pre-trained Transformer) series, that brought about significant breakthroughs in LLM technology.

With GPT models, the pre-training phase involves exposing the model to a large corpus of text, allowing it to learn language patterns and build a contextual understanding of words and phrases. The fine-tuning phase then refines the model's capabilities for specific tasks or domains, making it adaptable and versatile for various applications.

The practical applications of LLMs in business operations are vast. LLMs can assist in automating repetitive tasks, generating personalized content, and analyzing customer feedback. Additionally—and of most interest to the audience of this book—LLMs can process vast amounts of text documents, extract relevant information, and summarize key findings, helping to extract insights from large volumes of unstructured data. This capability can assist in research and development, data analysis, and decision-making processes, enabling businesses to derive insights from diverse sources of information more effectively.

For example, an LLM-based system can analyze maintenance reports, sensor logs, and operator notes to help operators efficiently navigate and discover relevant data, leading to better decision-making and improved operational efficiency.

LLMs can also play a vital role in industrial data analysis by assisting in critical activities such as anomaly detection and quality control. By ingesting historical data, sensor readings, and operational parameters, LLMs can learn to identify early signs of equipment failure, detect deviations from normal operating conditions, or pinpoint potential quality issues, supporting proactive maintenance strategies.

LLMs are a powerful tool for industry, improving operations in various ways that minimize downtime, reduce costs, and achieve higher overall efficiencies. With their ease of use, adaptability, and practical applications, LLMs offer a user-friendly solution that can streamline operations, automate tasks, gain valuable insights, and drive innovation in their respective industries.

Defining Industrial Knowledge Graphs



Extract Data Relationships, Capture Interconnections, and Trace Data Lifecycles

Knowledge graphs are constructed by combining data sets from diverse sources, each varying in structure. The harmony between schemas, identities, and context contributes to the coherence of this comprehensive data repository.

Schemas establish the fundamental framework upon which the knowledge graph is built, while identities efficiently categorize the underlying nodes. Context, on the other hand, plays a pivotal role in determining the specific setting in which each piece of knowledge thrives within the graph.



Knowledge graphs use machine learning to construct a holistic representation of nodes, edges, and labels through a process known as semantic enrichment. By applying this process during data ingestion, knowledge graphs can discern individual objects and comprehend the relationships between them. This accumulated knowledge is then compared and fused with other data sets that share relevance and similarity.

A knowledge graph enables question answering and search systems to retrieve comprehensive responses to specific queries. Knowledge graphs are powerful time-saving tools, streamlining manual data collection and integration efforts to bolster decision-making processes.

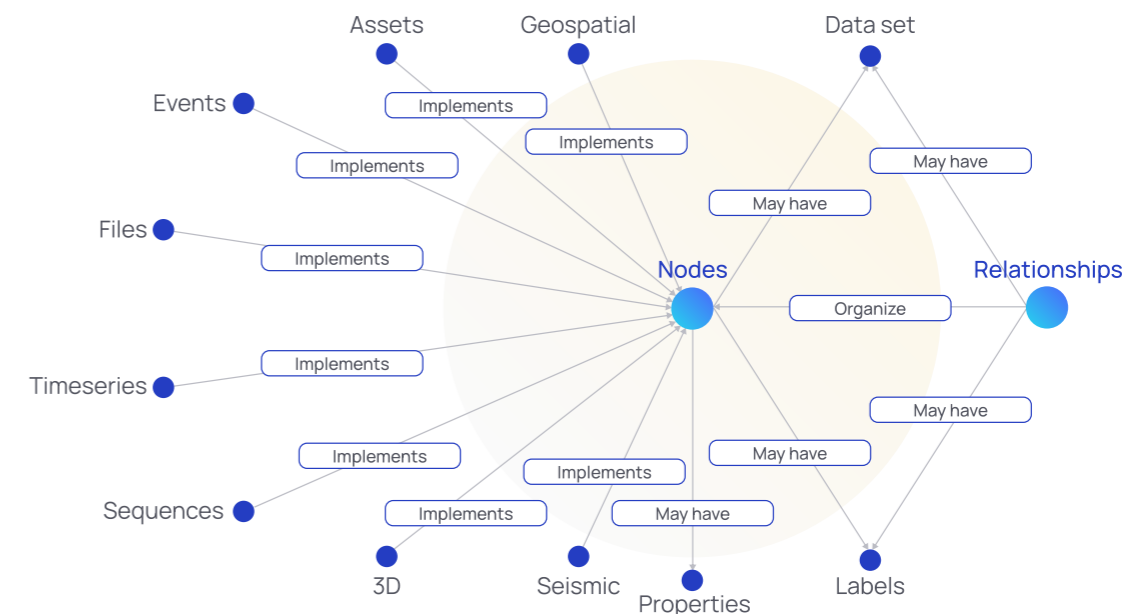
An industrial knowledge graph is an open, flexible, and labeled property graph data model that represents your operations.

It liberates the data that's been locked in different systems and applications (high-frequency time-series sensor data, knowledge hidden in documents, visual data streams, and even 3D and engineering data) and makes it meaningful and manageable.

An industrial knowledge graph turns raw data into valuable operational insights. With a knowledge graph, you can:

Go from search to discovery: When there are hundreds of data sources and countless name conventions, searching can be tedious. Discovering related data instantaneously will help you leap from antagonizing searches to discovering insights.

Understand relationships: And not just tables. The relationships often matter more to the answers we seek than the data entries they connect. Understanding our processes and systems requires understanding our data real estate in its context.





Execute industrial artificial intelligence (AI) initiatives: Knowledge graphs help operationalize your data science, artificial intelligence, machine learning (AI/ML), and Internet of Things (IoT) initiatives.

However, a knowledge graph is only as powerful as the data it can access. To be effective in complex industrial settings, a knowledge graph must include:

- Automated population with contextualization, cross-source-system IT, OT, and ET data
- A robust, well-documented API integration
- Extremely high performant, real-time, flexible data modeling to enable a variety of uses, including queries and search, natural language processing, machine learning algorithms, and visualization.

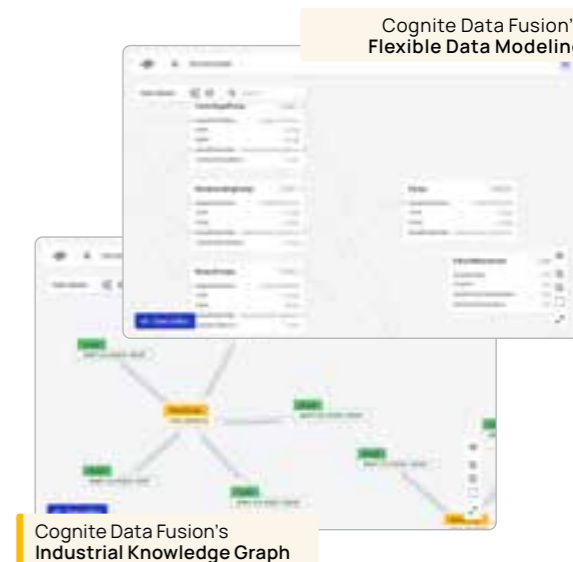
Like the principle of compounding interest, data in the industrial knowledge graph becomes increasingly more valuable as people use, leverage, and enrich that data.

More useful and high-quality data leads to more trusted insights. More trusted insights lead to higher levels of adoption by subject matter experts, operations and maintenance, and data science teams.

A user-friendly, AI-powered experience ensures adoption and use will grow, and this cycle repeats exponentially.



In this example, the left diagram above illustrates a simplified version of an industrial knowledge graph of a centrifugal pump. Depending on the persona, users may explore a problem with the pump from multiple entry points. Maintenance may start with the latest maintenance report, while an operator may use the time series, and a remote SME may begin with the engineering diagram (e.g., P&ID). The maintenance report, the work order, the time series values, and the engineering diagrams are each in separate systems. Having all this data connected in the industrial knowledge graph creates a seamless experience, regardless of the starting point.



This simple example illustrates the importance of data contextualization across different systems. With generative AI, Cognite's data contextualization capabilities power the industrial knowledge graph (as seen on the right side) to provide access to the maintenance report, work order, time series, and more in a single location.

With the industrial knowledge graph as the foundation, data is understood and structured to meet the specific needs of users or use cases. The topic of industrial knowledge graphs, data models, and digital twins is discussed in more detail in chapter seven.

Making Generative AI Work for Industry

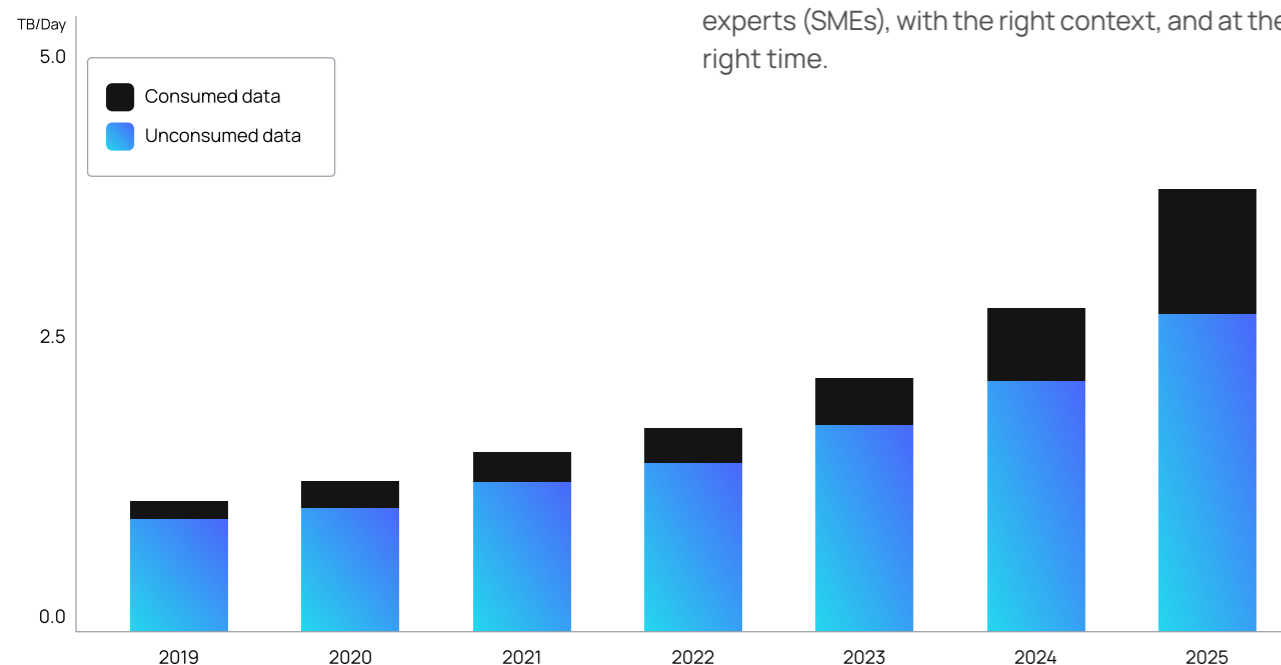


Liberating Industrial Data

To truly solve the industrial data and AI problem, data must be liberated from siloed source systems.

Data users still spend up to 80% of their time searching, gathering, and cleaning data, costing businesses millions of dollars in working hours every year. This bottleneck toward productivity will only become worse in legacy architectures as IDC predicts data generation in asset-heavy organizations to increase by 3X in the next two to three years.

Industrial organizations must first liberate all data across numerous siloed data sources, and then get the right data to the right subject matter experts (SMEs), with the right context, and at the right time.



Data has no value unless the business trusts and uses it.

Data contextualization also raises the question of trust. Industrial enterprises must trust the data they put into solutions, from dashboards to digital twins to generative AI-powered solutions. Ultimately, data has value only if the business trusts and uses it.

Organizations can increase trust and avoid other disadvantages of data lakes by adopting an Industrial DataOps mindset. This approach makes data more valuable to a growing audience of data consumers, both inside the enterprise and across its partner ecosystem.

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.

Industrial DataOps is about breaking down silos and optimizing the broad availability and usability of industrial data generated in asset-heavy industries including oil and gas and manufacturing.

SMEs, for example, must be empowered to access and harness data effectively. With simple access to complex industrial data, industrial organizations can bring together formerly isolated SMEs, departments, platforms, and data deployed by OT and IT teams to improve operational performance through unified goals and KPIs across the enterprise.



Why Historians are Not Enough

The source for most operational data is a data historian, which holds the data deemed valuable long term. But there is more to a plant than what is available in the historian. Because operators instrument the plant for safe operation and control, most of the data required for efficient maintenance must be picked up elsewhere, i.e., via inspections, maintenance campaigns, industry reports, vendor data, and more.

Additionally, what is available in the historian is limited by the data types supported. Modern instruments and protocols allow for more data types than numbers, mainly ignored or stored in separate, often local, databases.

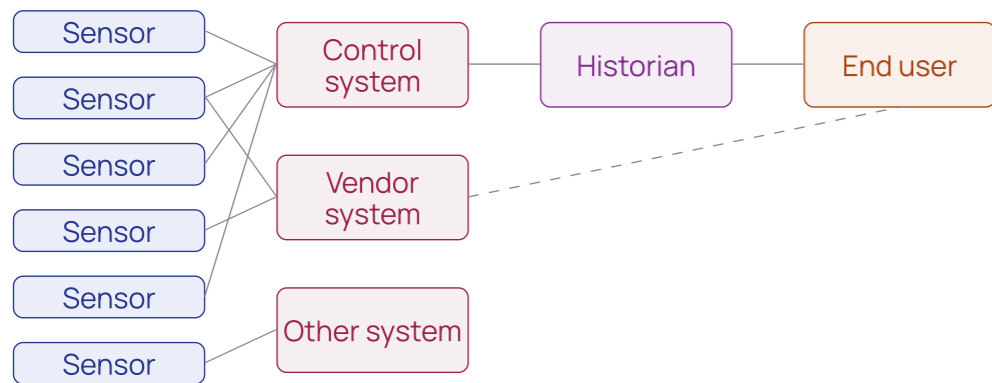
A modern pressure transmitter can provide functions, such as self-calibration, and housekeeping data, such as circuit board temperature, in addition to a pressure reading. If not planned for, these features go unused.

Also not in the historian are fast-sampled data stored as signatures, such as vibration or operational data, which are handled in specialized systems when the control system's input/output does not support it.

Data layers: simple data flow



Data layers: complex data flow for one type of data

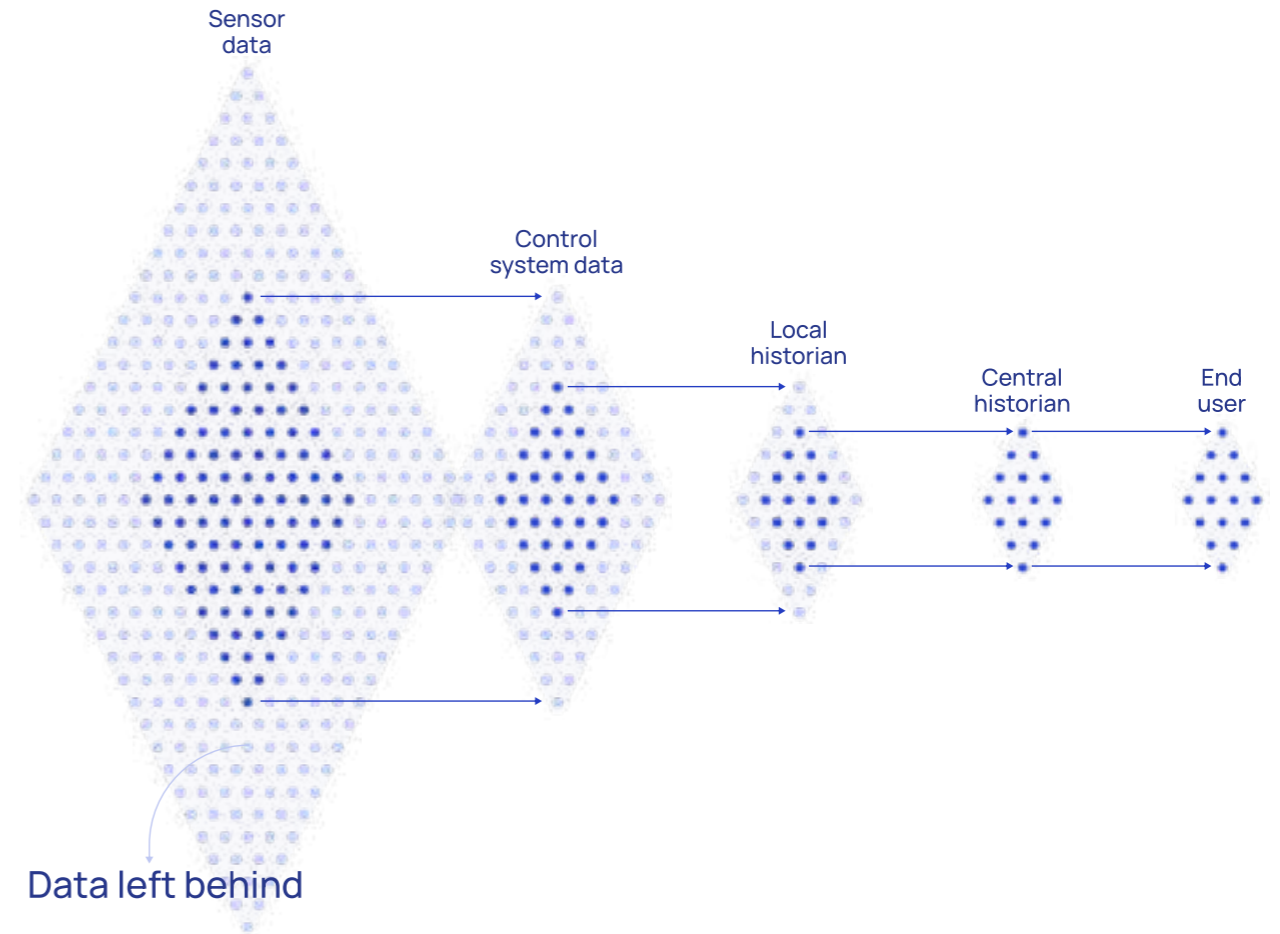


Only a fraction of the instrumentation data is consumed

Consumers are a mix of control systems, vendor solutions or on prem databases.

The historian is associated with process data from the control system. Limited by data type support and critically short term and long term.

End users access to data is limited to what is made available to the location where they are. Most often what is in the historian.



Too Much Data Left Behind

The problem with the current setup is that valuable data is left unused, either because it is filtered away or because it ends up in a local data store. There are several reasons why this happens:

- A new instrument with more capabilities replaces one with fewer, but the additional data is not hooked up.
- Including more data can be costly, as control systems and historians have a cost tied to I/O or tag count; however, some vendors have started to differentiate between types of data, e.g., process critical vs. housekeeping.
- Different vendors provide sensors, control systems, and historians, so changes are costly in terms of hardware and software and in terms of management and coordination with various vendors.
- There is a lack of knowledge or communication on how the additional data can add value to other contexts (e.g., the process operator prioritizes data from a sensor that meets specific process requirements but ignores built-in housekeeping data that can help to identify internal errors since it is not applicable for the primary use case).
- The operator designs instruments and control systems to last for many years. Because the focus is on safe operation and control, the people involved are rightfully conservative and hesitant to change something that works for its primary purpose.

An architecture and data flow seen above means that a large amount of instrumentation data—from housekeeping data to more complex data types—is left behind.

Even when organizations design architectures that pass this more detailed data along to the control system, it can be stopped by the next layer (i.e., the historian) since that solution vendor also charges for by I/O Count and has limitations in what data types it supports.

Today, sensors can self-calibrate and provide internal data about their health. So a pressure transmitter can report both the measured process pressure, as well as calibration curves, error registered, and onboard circuit health data—all through an ethernet interface and open protocol.

Though control system vendors and the heavy machine industry’s requirement for safety still require gathering the data locally and local control, control systems now depend more significantly on IT-like architectures, using ethernet and virtualization principles while still providing necessary safety and uptime requirements. The introduction of commoditized technology on this level enables control systems to handle more data types and serve them more efficiently for local users. It also opens up to separate the data pipelines for critical and non-critical data to reduce cost (less I/O and interface) and gain velocity (less interfaces and vendors).

While today’s historians capture the critical process data required for analysis, most do not extend their data types like engineering drawings (P&IDs, standard operating procedures, inspection rounds), work orders, reliability data, and IoT sensors.

But What About Hallucinations, Security, and Privacy?

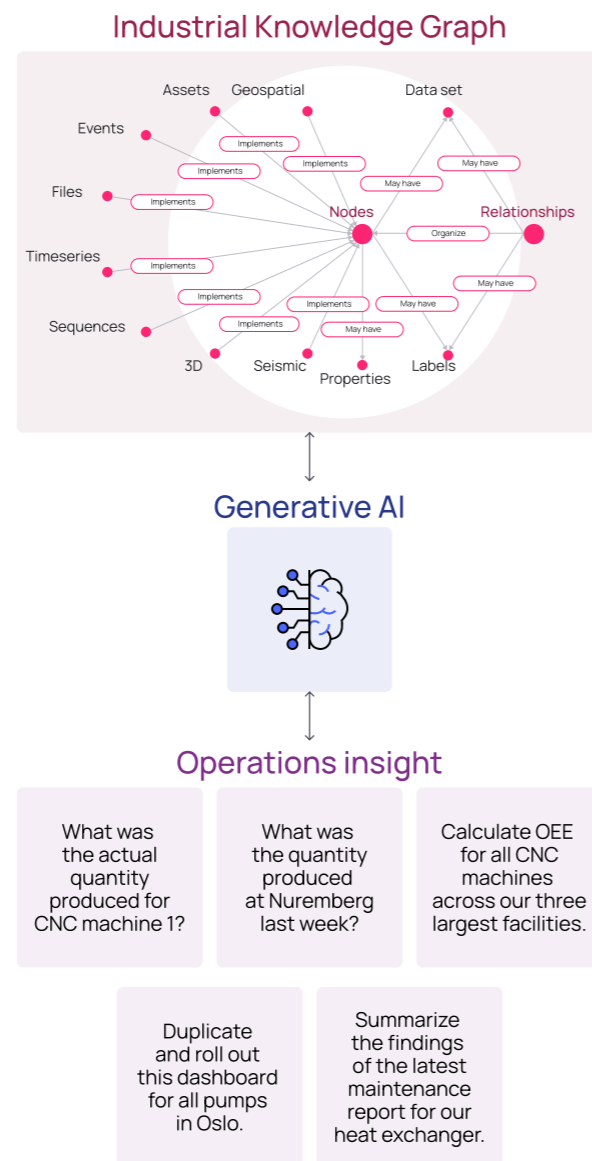
As powerful as LLMs are, there have been instances where they produce inaccurate or misleading information. These hallucinations can be problematic, especially when LLMs are used in critical decision-making processes.

LLMs have a vast knowledge base to draw from. However, the content in an LLM's data stores may be dated and based solely on content from the public domain. These factors limit the source data for generating a response, potentially leading to out-of-date info or 'creative' answers to compensate for the information gap (hallucinations).

If we can 'train' an LLM like ChatGPT on curated, contextualized industrial data, then we could talk to this data as easily as we converse with ChatGPT and have confidence in the basis of the response.

Contextualized data includes explicit semantic relationships within the data, which ensures that the text consumed is relevant to the task at hand. For example, when prompting an LLM to provide information about operating industrial assets, the data provided to the LLM should include the data and documents related to those assets and the explicit and implicit semantic relationships across different data types and sources.

The resulting industrial knowledge graph also processes data to improve quality through normalization, scaling, and augmentation for calculated or aggregated attributes. For generative AI, the adage of 'garbage in, garbage out' applies.



Aggregations of industrial data in large data warehouses and lakes that have not been contextualized or pre-processed lack the semantic relationships needed to 'understand' the data and the data quality necessary for LLMs to provide trustworthy, deterministic responses.

Security and privacy concerns also arise with LLMs, which can be vulnerable to adversarial attacks where malicious actors deliberately manipulate input prompts to generate misleading, harmful outputs or exploit vulnerabilities in the LLM implementation or infrastructure to gain unauthorized access to sensitive data or disrupt operations. For example, compromising an LLM used in a control system could lead to equipment malfunctions, production errors, or unauthorized access to critical resources.

Data leakage is a particular concern for industrial organizations that often deal with sensitive data, including proprietary designs, trade secrets, or customer information. LLMs trained on such data could unintentionally leak confidential information through the generated responses.

Moreover, there is a concern that fine-tuning LLMs on proprietary or confidential information may inadvertently expose sensitive details if the model is not adequately anonymized or protected.

Ensuring proper anonymization and protecting sensitive data during LLM training, protecting the LLM infrastructure, securing data transfers, and implementing robust authentication and authorization mechanisms are essential to mitigate cybersecurity risks and the risk of information leakage.

Overall, the challenges of hallucinations, security vulnerabilities, and privacy risks associated with LLMs highlight the need for rigorous evaluation of LLM-based systems. With that in mind, let us look at how Cognite addresses these concerns.

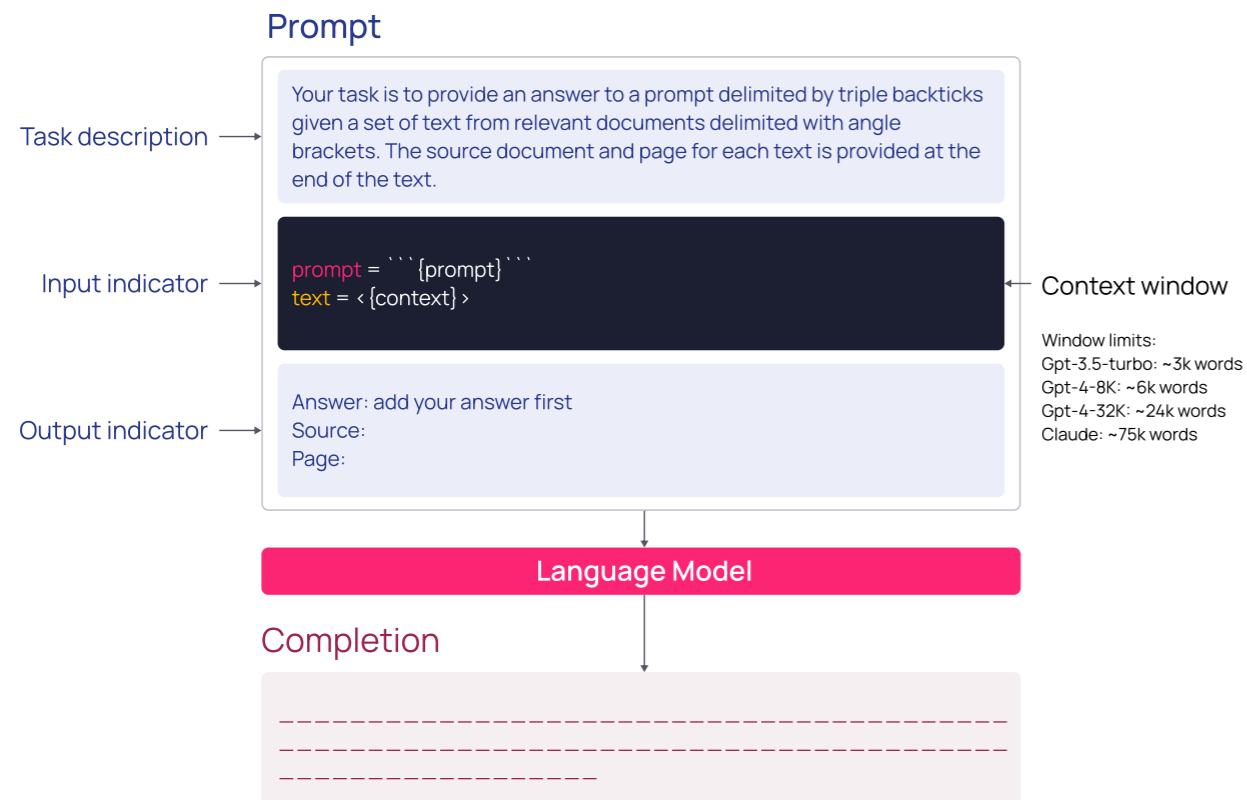
RAG is All the Rage

As discussed, LLMs have access to the corpus of text used during the model's training. However, as input, these models can also take new information to incorporate when responding to a natural language prompt. This additional content can come as real-time access to web-based queries of publicly accessible content or from the user in the form of additional inputs as part of the prompt.

Retrieval Augmented Generation (RAG) is a design pattern we can use with LLMs to provide industrial data directly to the LLM as specific content to use when formulating a response. A type of in-context learning, RAG lets us use off-the-shelf LLMs and control their behavior through conditioning on private contextualized data. This approach allows us to utilize the reasoning engine of LLMs to provide deterministic answers, based on the specific inputs we provide, rather than relying on the generative engine to create a probabilistic response based on existing public information.

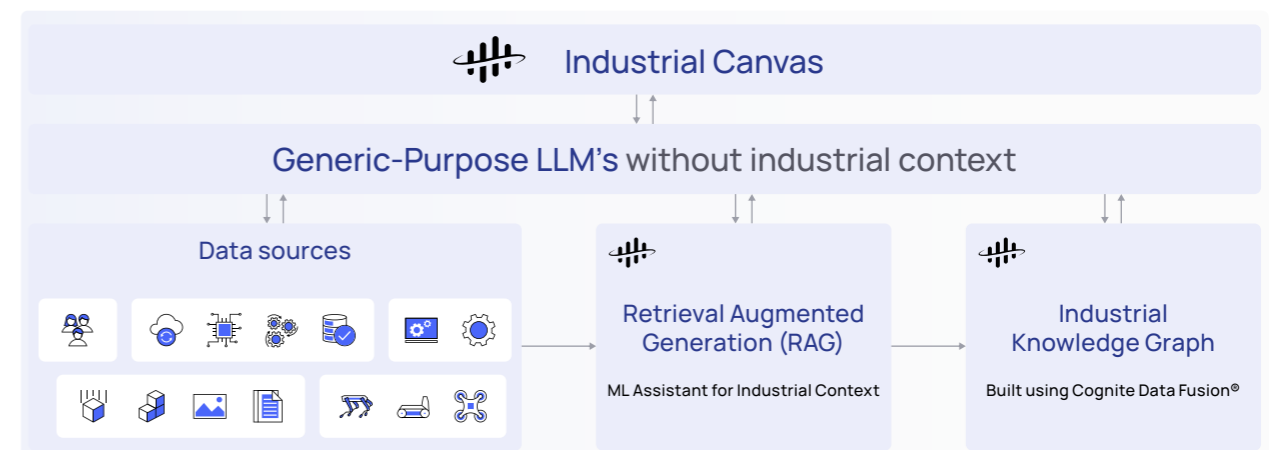
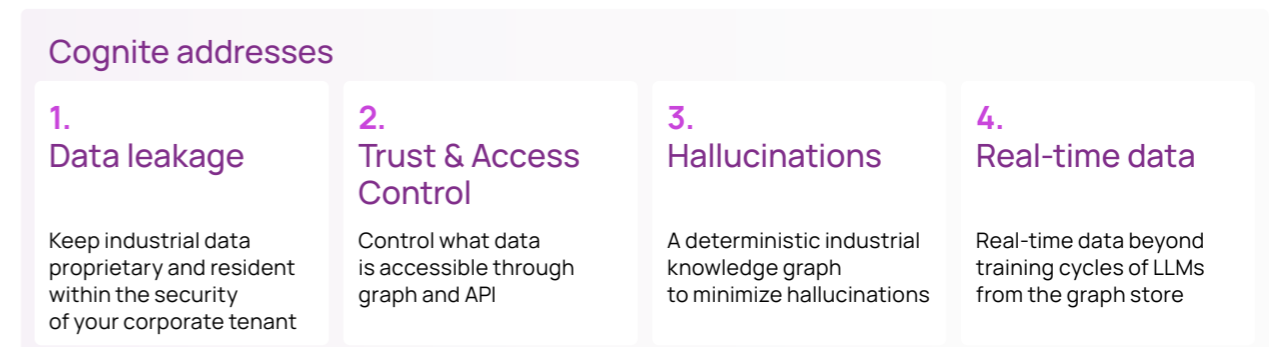
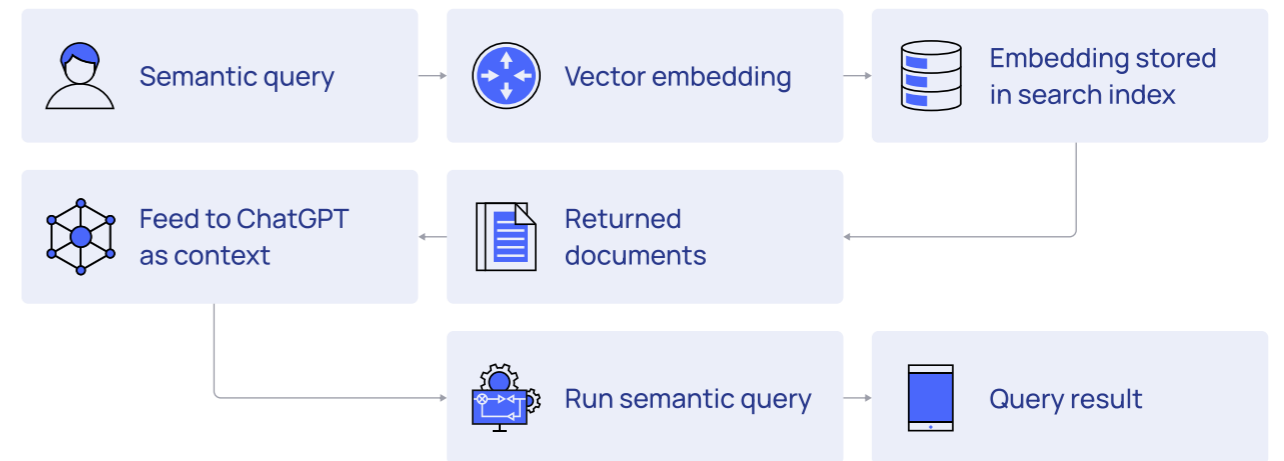
Another thing to note here is the context window limitations of LLMs. The context window is the range of tokens the model can consider when responding to prompts. GPT models start with a 2K window size (GPT-3) and go all the way to 32K (GPT-4). In this context, a token is a piece of text that could be as short as one character or as long as one word. GPT-3's context window can handle about 2048 tokens (nowhere near large enough to take input from an entire industrial database).

By contextualizing proprietary industrial data to create an industrial knowledge graph, we can convert that enriched content into embeddings and store it in a private database, fine-tuned for embedding storage and search (vector database).



Codifying this context as an industrial knowledge graph is critical to leveraging your vast industrial data within the context window limitations and enabling consistent, deterministic navigation of these meaningful relationships. By utilizing the open APIs of major LLMs, we can then leverage this trusted source of industrial context to create and store embeddings in a way that becomes searchable (semantically) and enables us (with minor prompt fine-tuning) to fully leverage the reasoning engine of LLMs to give us actionable insights.

Leveraging this pattern, we can keep industrial data proprietary and resident within the security of your corporate tenant. We can maintain and leverage the access controls required to maintain large enterprises' trust, security, and audit requirements. Most importantly, we can get deterministic answers to natural language prompts by explicitly providing the inputs LLMs should use to formulate a response.



A Thought on Cost

The clever readers might wonder, “Won’t the context windows supported by foundation models quickly become bigger, making RAG obsolete?” The short answer is no.

The longer answer is cost.

In all algorithms, even linear scaling—the best theoretical outcome possible—is simply a deal breaker on cost for almost all use cases imaginable. At current API rates, each query to GPT4 would easily end up costing tens of dollars at a context window still limited to below 1,000 pages (which is hardly any enterprise’s data corpus).

When thinking of these queries similar to Google searches by your employees, it quickly becomes apparent why context window optimization—and, of course, hallucination mitigation—using RAG as a cost-reduction initiative is already as hot as LLMs themselves.

The significant difference in the cost between sending queries to GPT3.5 or GPT4 alone will force AI application developers to carefully optimize cost (known) vs. value (known for dependence on quality, unknown for ultimate business value).

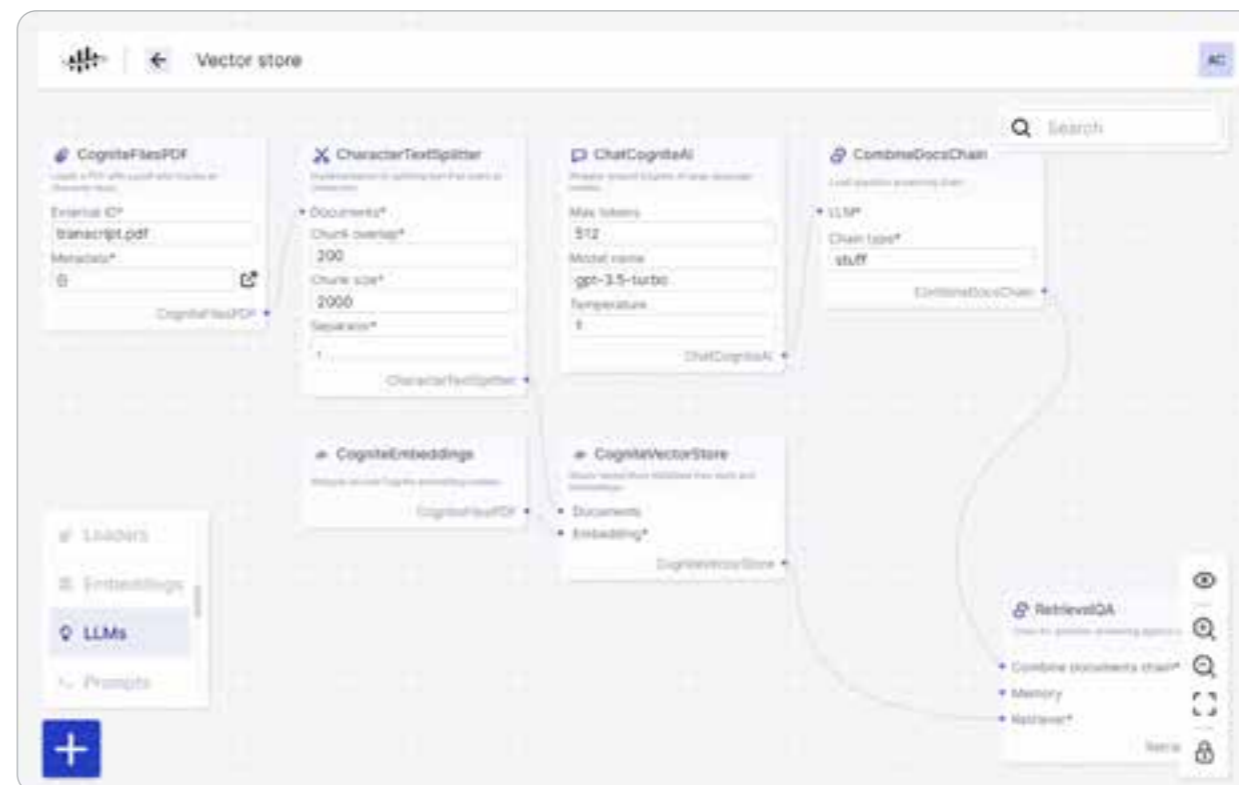
Model	Completion	Prompt	Context	TPM	RPM	1M tokens
<code>gpt-4-0314</code>	\$0.06/1K	\$0.03/1K	8192	40000	200	\$1200.00
<code>gpt-4-32k-0314</code>	\$0.06/1K	\$0.12/1K	32768	80000	400	\$1200.00
<code>gpt-3.5-turbo</code>	\$0.002/1K	\$0.002/1K	4096	40000	200	\$4.00

In LLM application development, choosing the best each time will come at a very high cost to those using it. One could argue that RAG is to LLMs what data engineering is to analytics: carefully selecting and preparing the correct data before sending it for analysis.

Back to the need for rigorous evaluation of your AI systems, look for tools with an easy-to-use UI that enables experimenting and prototyping data flows. The example on the next page is a no-code UI utilizing vector store, embeddings, and LLM while reading data from an industrial knowledge graph.

Here we see that we can embed PDF documents from the industrial knowledge graph and, with context from these documents, allow the LLM to perform a semantic search.

Behind the scenes, this particular flow is a standard flow for Document Q&A. The difference is that the UI makes the underlying components more visible and helps explain the flow visually. Visual data flow preparation can extend to more complex flows and help developers optimize AI applications and reduce the overall, long-term costs of AI infrastructure management.



AI Agent Frameworks

RAG is just one component of an industrial AI architecture. While RAG effectively solves hallucination and data-freshness problems, AI Agent frameworks give AI applications new capabilities.

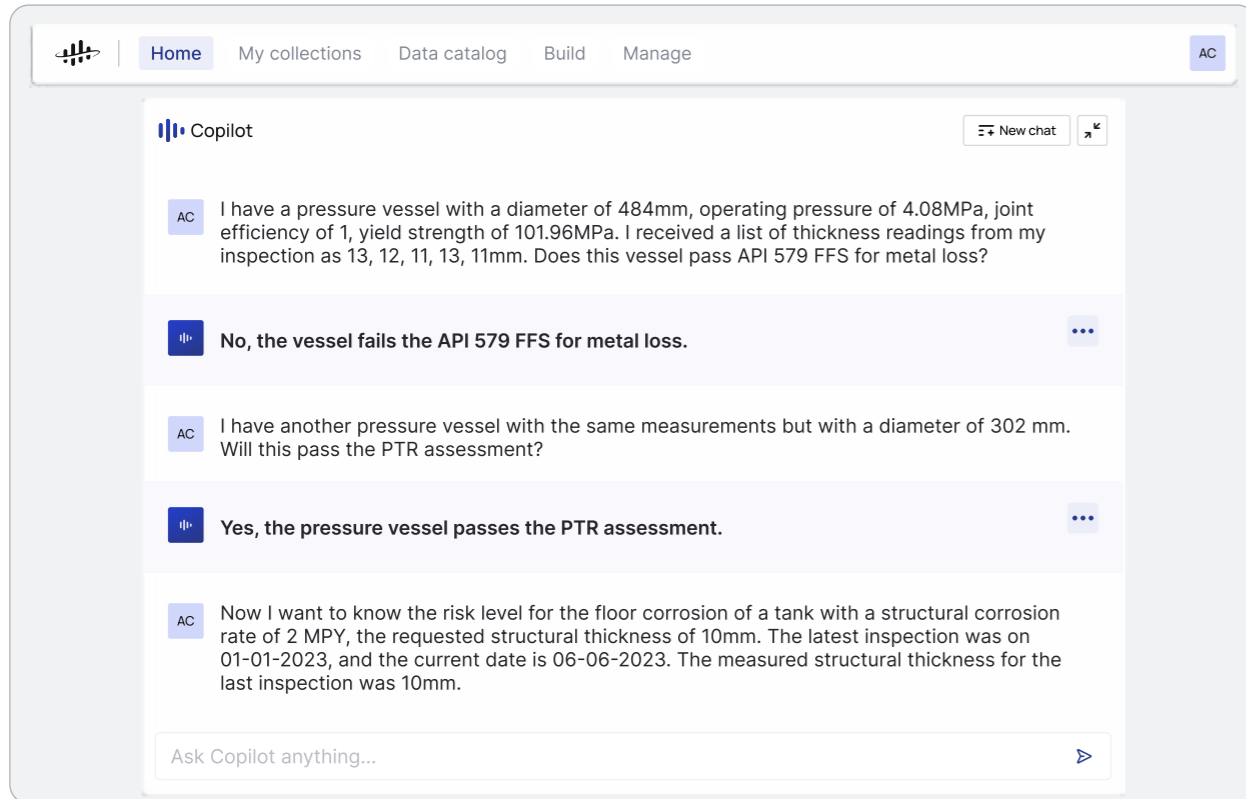
Agents are designed to achieve specific goals and can perceive their environment and make decisions autonomously. Agents include chatbots, smart home devices and applications, and the programmatic trading software used in finance.

Agents are classified into different types based on their characteristics:

- **Reactive agents** respond to immediate environmental stimuli and take actions based on those stimuli.
- **Proactive agents** take initiative and plan to achieve their goals.

- **Fixed environments** have a static set of rules that do not change.
- **Dynamic environments** are constantly changing and require agents to adapt to new situations.
- **Multi-agent systems** involve multiple agents working together to achieve a common goal, often requiring coordination and communication to achieve their objectives.

Agents are another centerpiece of industrial AI architecture. And existing frameworks such as LangChain have incorporated some agent concepts already that can be incorporated into industrial solutions.



For example, Cognite has prototyped an AI copilot for reliability-centered maintenance using LangChain technology to better equip operators and reliability engineers to check damaged equipment. The copilot incorporates standards, documentation, and images to run high-fidelity engineering calculations through a human-language interface.

The value of this approach is that it combines the power of running complex mathematical calculations with the easy interface of a language model without compromising accuracy, which is a challenge in using LLMs.

However, it is important to note that copilot- and AI agent-based approaches leverage the power of natural language to understand and write code based on published API documentation and examples. This is impossible with data lakes or data warehouses where, without a contextualized industrial knowledge graph, there are no API libraries that can be used as a reliable mechanism to access rich industrial data. Additionally, because all data access happens through the APIs, no proprietary data is shared with third parties, and the built-in mechanisms for logging and access control remain intact.



The Industrial Data and AI Problem



The State of Data Liberation and Data Contextualization

For every one person who can 'speak code,' there are hundreds of others who do not, especially in the industrial environments where there are numerous data types and source system complexity. To use industrial data broadly, it requires context.

Subject matter experts (SMEs), field engineers, and data scientists deserve simple access to all industrial data in a single workspace. This requires a unique way to leverage and apply contextualized data. Industrial applications, as we know them today—especially data dashboarding and visualization—will soon be 100% transformed.

Open, composable workspaces with integrated AI copilots will become the point-of-entry to engage with industrial data in the same way the web browser replaced desktop applications. The infusion of generative artificial intelligence (AI) accelerates the ability to converse with contextualized, trusted data via a performant API—without writing a single line of code.

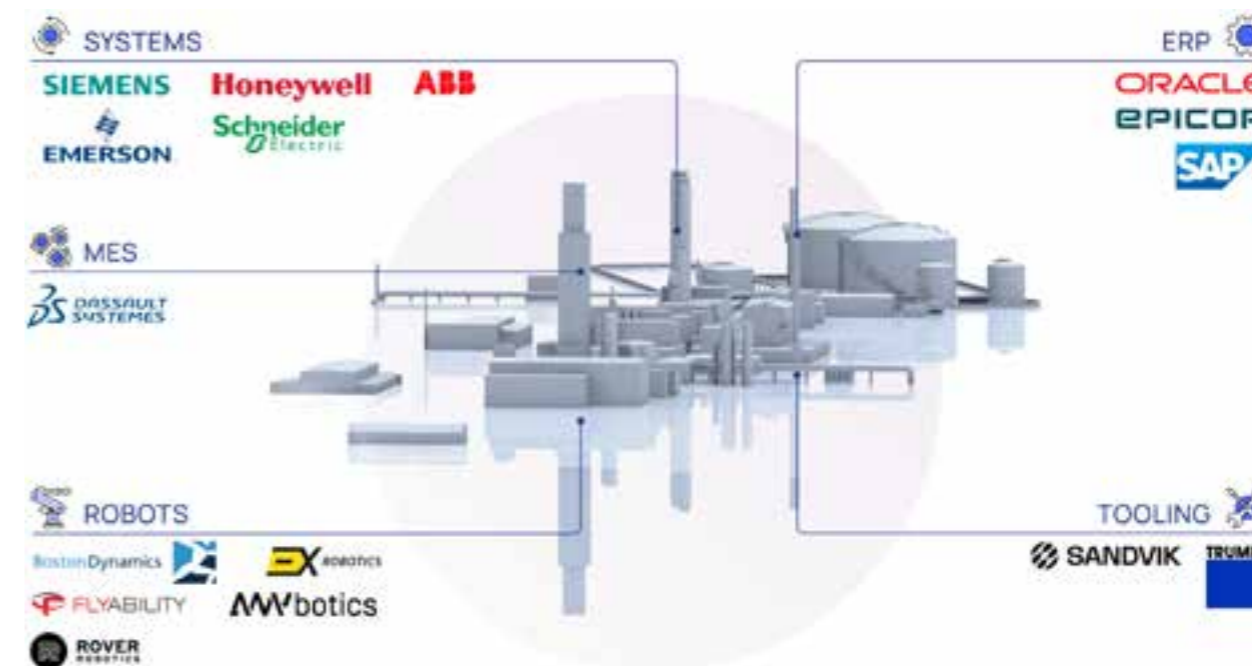
Generative AI also thrives on context. While generative AI has tremendous potential, answers are often wrong without data that is contextualized in an industrial knowledge graph. LLMs such as ChatGPT are trained on 10-100s of billions of parameters, yet data sets of this size don't exist for industrial environments. If a pre-trained model is integrated with raw data in a data lake, the patterns aren't readily identifiable to the model. Using generic LLMs on uncontextualized, unstructured industrial data significantly increases the risk of hallucinations.

To enable generative AI solutions to provide the right answers in industrial environments, there must be an efficient way to provide generative AI solutions with more context.

Drowning in Data, Starving for Context

A typical industrial facility can have 100,000 data points continuously updating, some faster than once per second, across more than 50 applications.

Unstructured data types are also increasing with more reliance on images, videos, acoustic, 3D models, pointclouds, and engineering drawings to provide additional context on the state of operations. Traditional efforts to connect data from systems are manual, time-consuming, and fail to manage structured data at scale, much less incorporate the growing unstructured data.

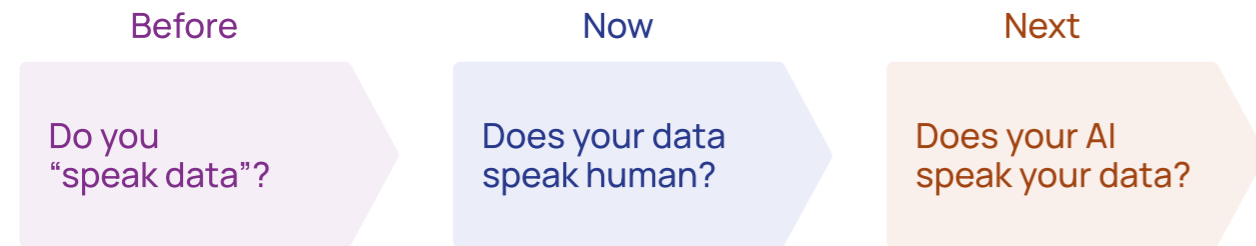


Enterprises large and small are rushing to reduce the barriers their workforces must overcome to consume data—or to be more data literate.

Gartner formally defines data literacy as, "the ability to read, write and communicate data in context," more informally expressed as, "Do you speak data?" Data literacy includes understanding data sources and constructs, analytical methods and techniques applied to data, and the ability to describe the use case applications and resulting value.

The need to understand industrial data is also relevant to the growing demand to apply generative AI in these environments, where large language models (LLMs) like ChatGPT lack the industrial context required to provide deterministic, trusted responses unique to each facility.





In the past, people were forced to 'speak data' to gain actionable industrial insight. Now, we have progressed to make industrial data speak human. We can liberate complex industrial data and put it into context for the people and equipment who need it.

Similarly, AI investments must speak to your people about your industrial data. Those who can already provide simple access to complex industrial data deliver more business value with an ROI above 400% and are well-positioned to use large language models to unlock new business opportunities rapidly.

While the urgency for subject matter experts (SMEs) to become more data literate is clear, they face the industrial data and AI problem. The physical world of industrial data is a messy place. For example, equipment wear, fluctuating operational targets, and work orders are all important when assessing the root cause of an operational issue. A process variable cannot be observed and interpreted in isolation. Is the variable trending as expected? What is the maximum value recommended by the vendor? When was this equipment last inspected, and what were the observations?

These questions concern the SMEs and field engineers responsible for keeping equipment running and continuously operating at a level that optimizes short-term and long-term efficient production. Increasingly, documentation and electronic trails of maintenance and operational history can be accessed digitally.

However, diagrams and vendor documentation are still found in static PDF files, maintenance records are scanned paper documents, and 3D models are not up to date with the physical world as modifications and maintenance have occurred over the asset's lifetime. How do we enable SMEs across operations, maintenance, and quality to solve their problems with digital tools?

To adequately empower SMEs to extract the value of industrial data, operationalizing data must become a core part of your business strategy. Data must be liberated, contextualized, and easily used to run accurate AI models and generate data-driven insights. A robust data foundation powered by a strong contextualization engine is crucial to optimize production and enhance operational efficiency.

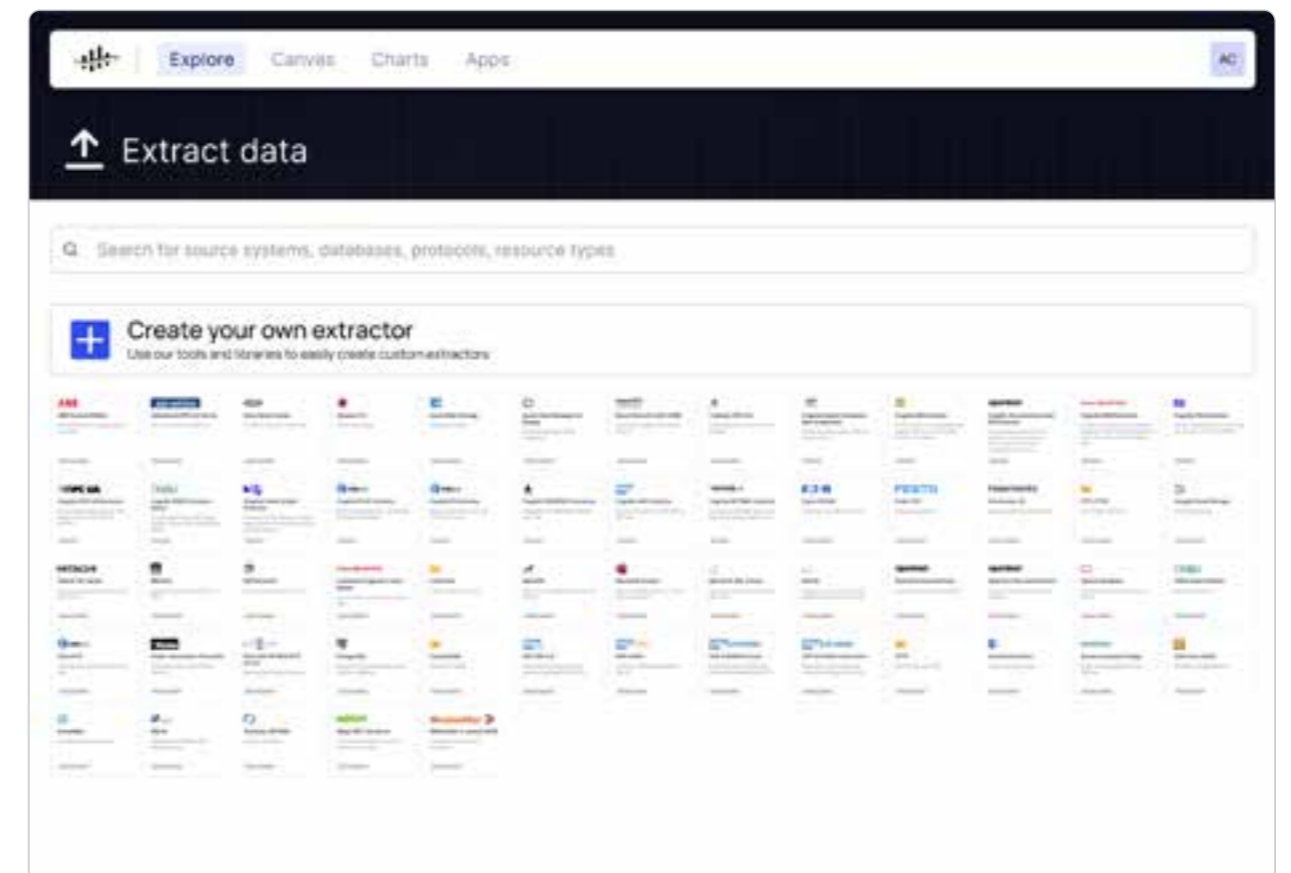
The foundation will enable operational teams to better work with data and improve their work efficiency. It will also become a solid foundation to run generative AI models that accelerate workflows, optimize production, automate repetitive tasks, assist in decision-making, and more.

A Modern Approach to Liberating Industrial Data

Reducing the burden of liberating industrial data requires an edge-to-cloud approach to integrate IT, OT, or engineering data for prioritized use cases. The edge solution must be able to collect data from legacy controllers, siloed equipment skids, IoT sensors, and more to address the fragmented OT systems.

The cloud solution must be able to integrate this edge data with existing OT data sources (historians, MES, SCADA, quality), IT sources (ERP), and engineer data (3D, point clouds, P&IDs, etc.).

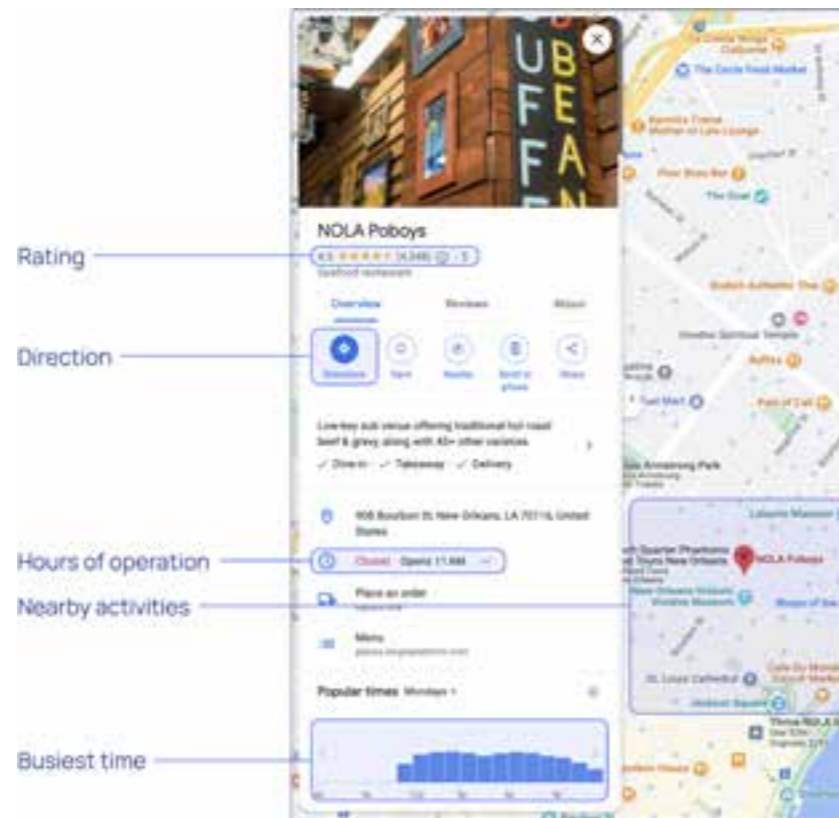
Most importantly, liberating data from existing industrial sources must not require a custom extractor development for each new source. Addressing this issue requires pre-built extractors into the most common industrial data sources. Connecting to new source systems must take a maximum of hours, not months, and the configuration must enable users to schedule frequency (from real-time streaming to daily updates) with full visibility into the state and performance of the connections. The connections must be stable, fully supported, and able to be monitored by an owner in the event of a disruption. All of these capabilities are prerequisites for having trusted data that is available for contextualization.



It's All About Relationships

Contextualization of data means to uncover and identify relationships between elements of data that are connected, but where the relationships are not explicitly represented. Identifying these connections can be done in many ways depending on the type of data and the kind of

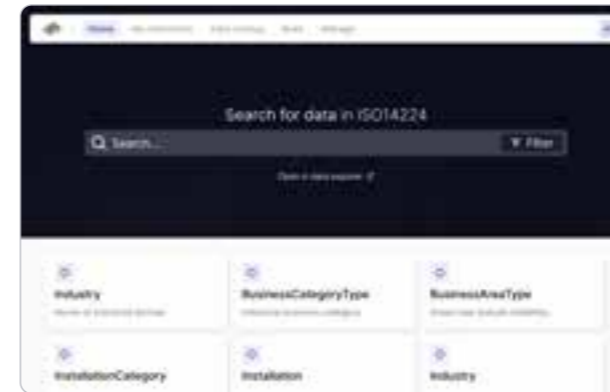
relationship—some of the techniques are simple pattern matching, others can rely on being able to understand specific data formats, having a lot of domain knowledge, and an ability to identify patterns that are not exact matches. The value of data contextualization is the automated discovery of real-time information.



Take Google Maps, for instance, which effortlessly combines map data with information from the web (rating, direction, hours of operation, nearby activities, busiest times) to provide users with a unified interface to quickly answer multiple questions from a single, seamless experience. In many cases, Google Maps often anticipates user queries even when incomplete and provides helpful

information and quick answers to questions. The user can then navigate the data in context to take an informed action.

But what about upstream energy production, downstream and process manufacturing, hybrid and discrete manufacturing, power generation, and other industrial environments?



Given the high-stakes nature of these industries and the sheer amount of operational questions being asked, operators in these environments deserve to easily navigate all related data in a



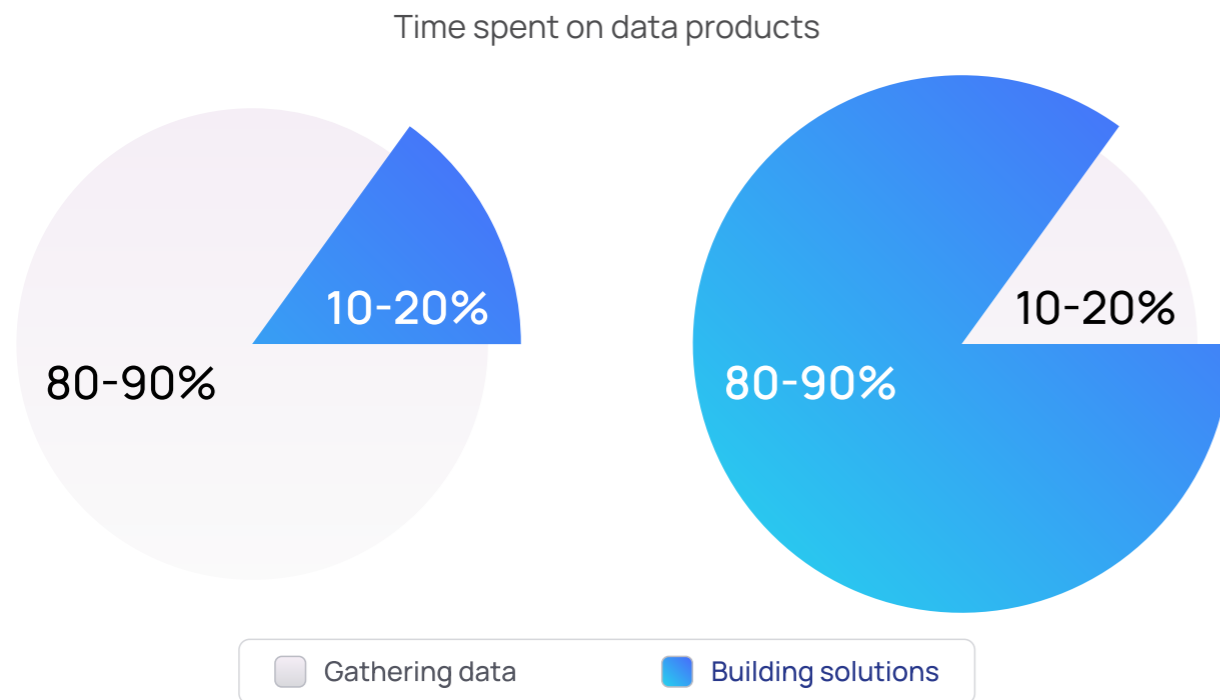
Google-like search, enhanced by generative AI, to make faster decisions, ensure safer working conditions, and improve asset reliability and resiliency of operations.



Although consumer and industrial technologies are different, the approach is quite similar. Like Google Maps, industrial solutions must put structured and unstructured data, real-time and historical, into business context, enabling anyone to build, deploy, and scale digital solutions



that drive business value. To answer operational questions, users can interact with contextualized information with a Google-like search, a 3D view, or a drawing (P&IDs or process flow diagrams) and, through any starting point, quickly arrive at the information they need.

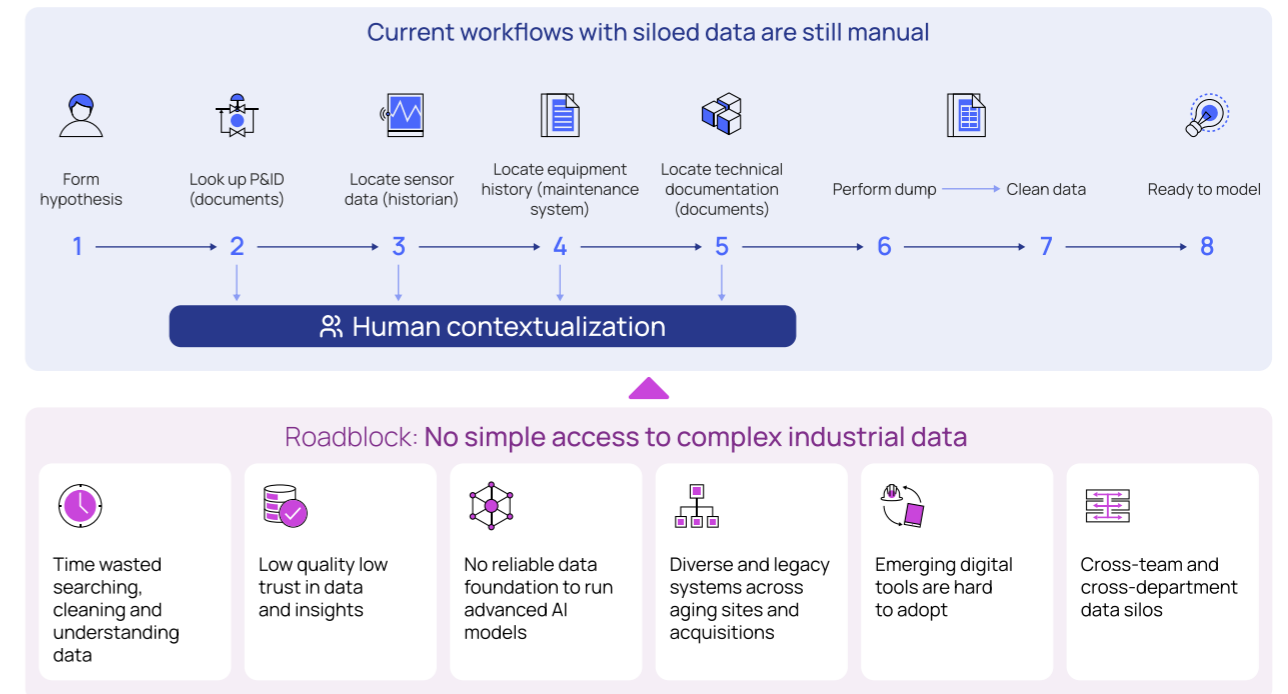


Working with contextualized data, the day-to-day decisions of the user become faster and more data-driven.

Instead of SMEs spending 80% of time finding and aggregating information, contextualized data flips the script and empowers end users to spend less time gathering data and more time focusing on making better-informed decisions with the help of generative AI and copilots.



Industrial data, in its siloed form, fails to unlock today's business opportunities



Unfortunately, today's reality for industrial search is too complex. SMEs spend hours looking for and interpreting data they need to tell an asset-centric narrative only to learn that the data is incomplete, inaccurate, or that the decision no longer matters. This time is spent looking through reports and spreadsheets, engaging with other data owners, or making new requests to IT—none of this work happens through a single pane of glass.

While many previous attempts have been made to solve the industrial data and AI problem using solutions such as data lakes, they often fail. These data solutions present disadvantages, including chaotic data governance, privacy issues, and inability to integrate data and share data changes that can delay time to value, increase the gap between SMEs and IT teams,

and increase cybersecurity risks and costs. Even where data sources are connected within a data lake, data often needs more context due to limited documentation or information loss due to inconsistent structure or tagging. As such, data lake solutions cannot establish sufficient relationships between all data in the lake, making it usable by only a few experts (not SMEs) who know how to navigate the data lake effectively and increasing the likelihood of hallucinations from generative AI solutions.

To truly solve the industrial data and AI problems, data must be liberated from siloed source systems and put into context so it can be used to optimize production, improve our asset performance, and enable AI-powered business decisions.

A Robust Contextualization Engine



Data Contextualization as a Foundation for Innovation

Contextualization means establishing meaningful relationships between data sources and types to traverse and find data through a digital representation of the assets and processes that exist in the physical world.

Continually contextualizing disparate data sources is an iterative process that creates a rich data foundation for operational innovation.

As relationships between previously siloed data sources are established in this data foundation, you naturally start building an industrial knowledge graph tailored to your operations.

The knowledge graph continuously evolves and spans many dimensions and data types, from the time series values to diagrams showing the process flows to a point cloud 3D model with recent images from an inspection.



Data from the industrial knowledge graph is deterministic, trustworthy, and accessible through a robust and performant API. These characteristics make the industrial knowledge graph necessary for enabling generative artificial intelligence (AI) use cases, modeling data for open industrial digital

twin use cases, and providing an open workspace to generate insights with contextualized data (Cognite Data Fusion's Industrial Canvas). Each of these topics will be discussed in detail, but first, we will start with how to contextualize data.

Industrial Canvas

Your teams deserve the industrial data & analytics UX only available with Industrial Canvas

- Simple access to complex industrial data in one collaborative workspace
- Use AI to develop and analyze complex cross data-source scenarios 90% faster
- Troubleshoot with real-time data and collaborate on RCA across teams and even sites

Open Industrial Digital Twins & Industrial knowledge graph

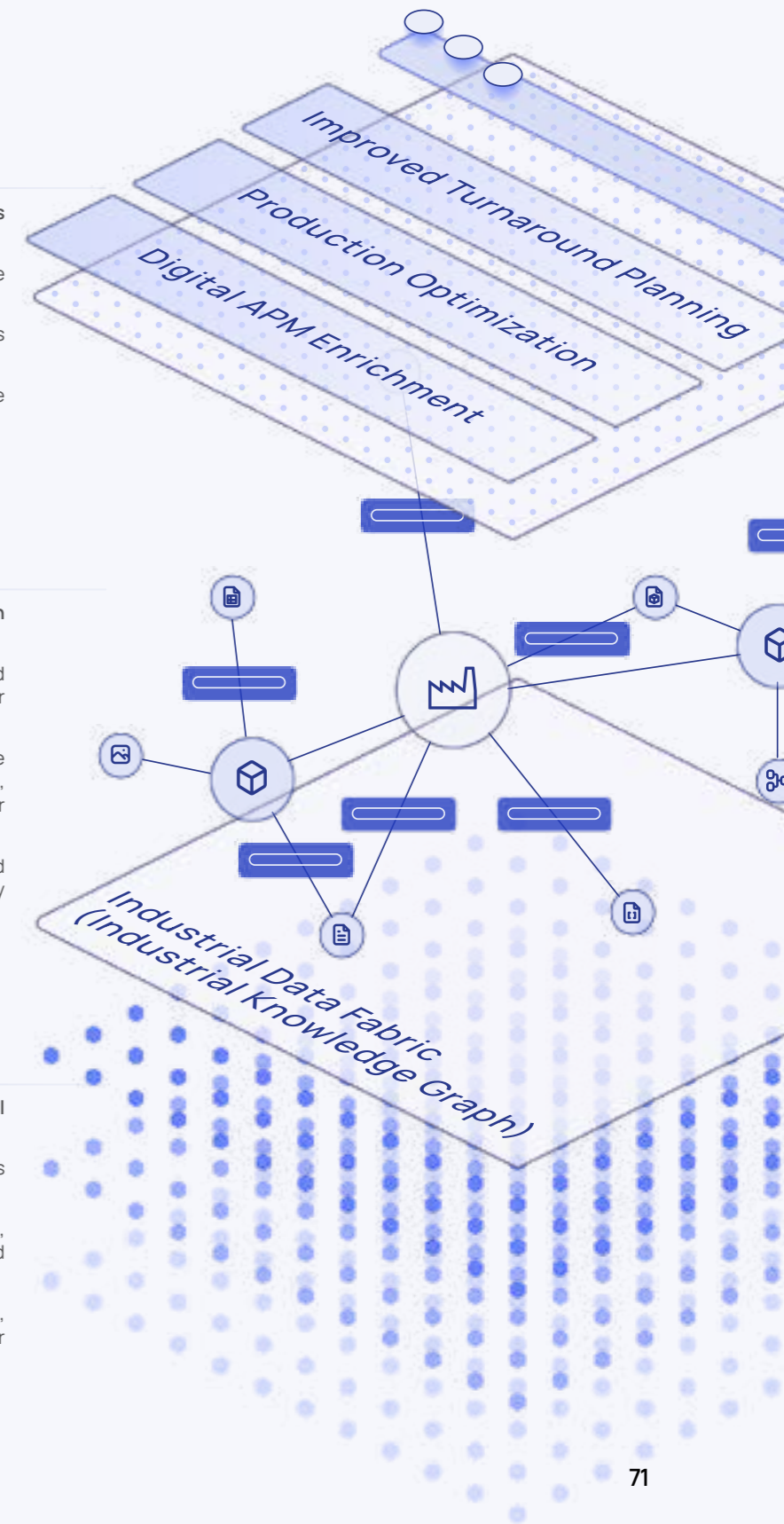
Represent our full enterprise as data graph with Open Industrial Digital Twins

- Build dynamic models of physical assets and processes, delivering real-time intelligence for active decision making
- Use data product practices to provide simple access with strong governance across source, domain, and solution data models across your ecosystem
- Take advantage of LLMs with Retrieval-Augmented Generation (RAG) powered by your own proprietary data, without data leakage

Scalable Industrial DataOps foundation

Supercharge your data engineering with Industrial DataOps

- 100+ OT, ET, IT, Simulator, and Robotics connectors out-of-the-box
- Connect from edge to cloud with confidence, with granular access controls, data lineage, and quality assurance built-in
- Catalog and activate you data across source, domain, and solution data models across your enterprise



Building Contextualization Pipelines

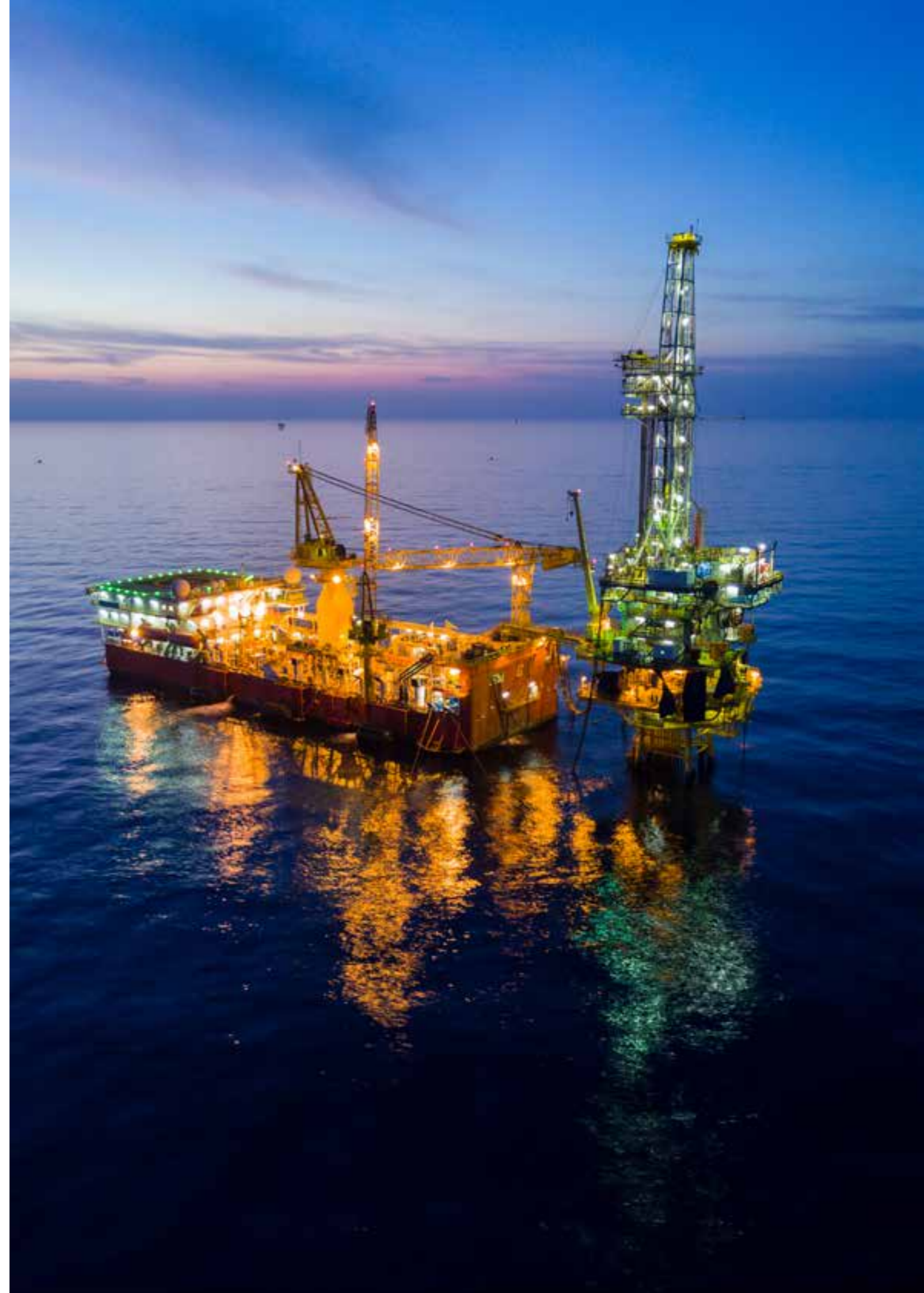
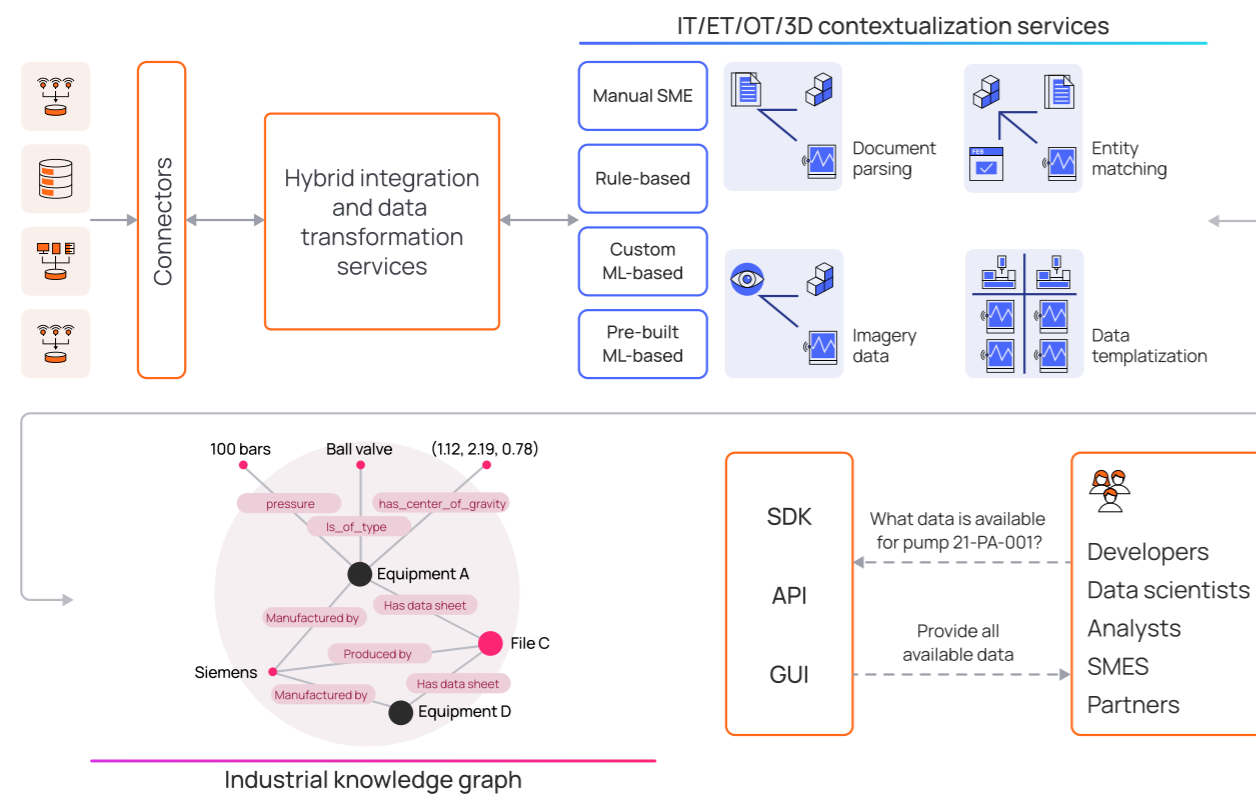
Data contextualization involves connecting all the data to better understand an asset or facility. Data relationships must be created through contextualization pipelines that develop and maintain a comprehensive, dynamic industrial knowledge graph.

Using contextualization pipelines addresses two key factors:

- Manual attempts to contextualize data are time-consuming upfront, whereas contextualizing data requires thousands of upfront hours to complete.

- Manually contextualized data is impractical to maintain across hundreds of thousands of mappings, and system changes are only captured through more manual efforts.

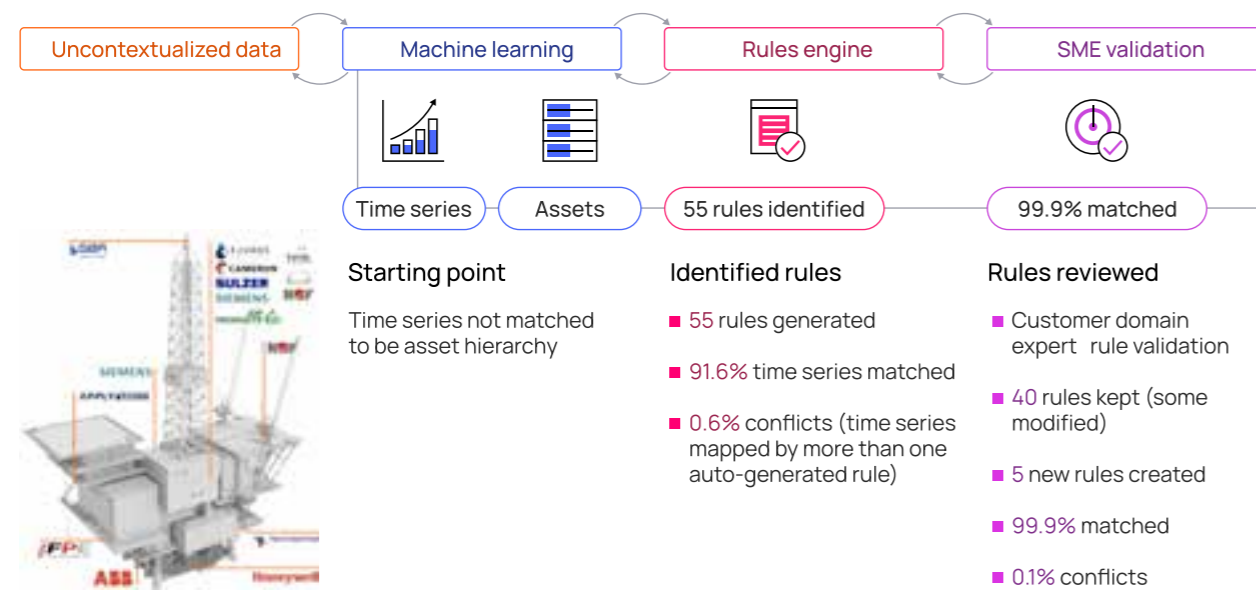
Contextualization pipelines must be able to connect all OT, IT, and engineering data types for a clearer understanding of an asset or facility and to reuse the industrial knowledge graph across many business solutions and business domains.



Achieving contextualized data at scale within a single site and, more broadly, across an enterprise requires establishing many contextualization pipelines through AI-powered contextualization

services that use pre-trained ML-based models, custom ML-based models, a rules engine, and manual/expert-sourced mappings.

Examples of contextualization pipelines



AI-powered contextualization permeates industrial data management, shifting the emphasis from data storage and cataloging to a true human data discovery experience, assisted further by a generative AI copilot.

The complete scope of contextualization capabilities to capture all structured, unstructured, and semi-structured industrial data extends to the following types of contextualization:

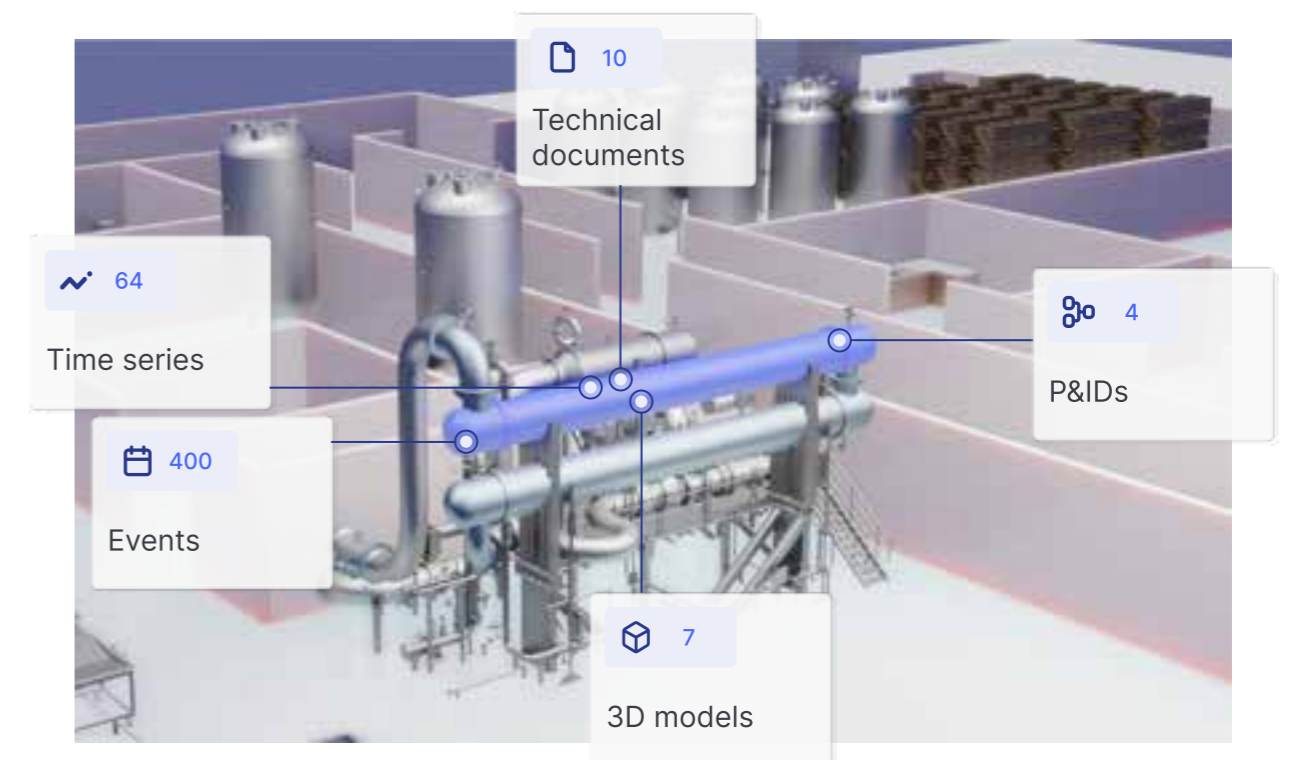
- **Entity matching** – map and connect time series, events, and tabular data to assets.
- **Create interactive diagrams** – map and connect tags on engineering diagrams (P&IDs or PFDs) to other data sources.
- **Contextualize imagery data** – upload and extract information from imagery data.
- **Create document classifiers** – label training data and create document classifiers.

Entity Matching

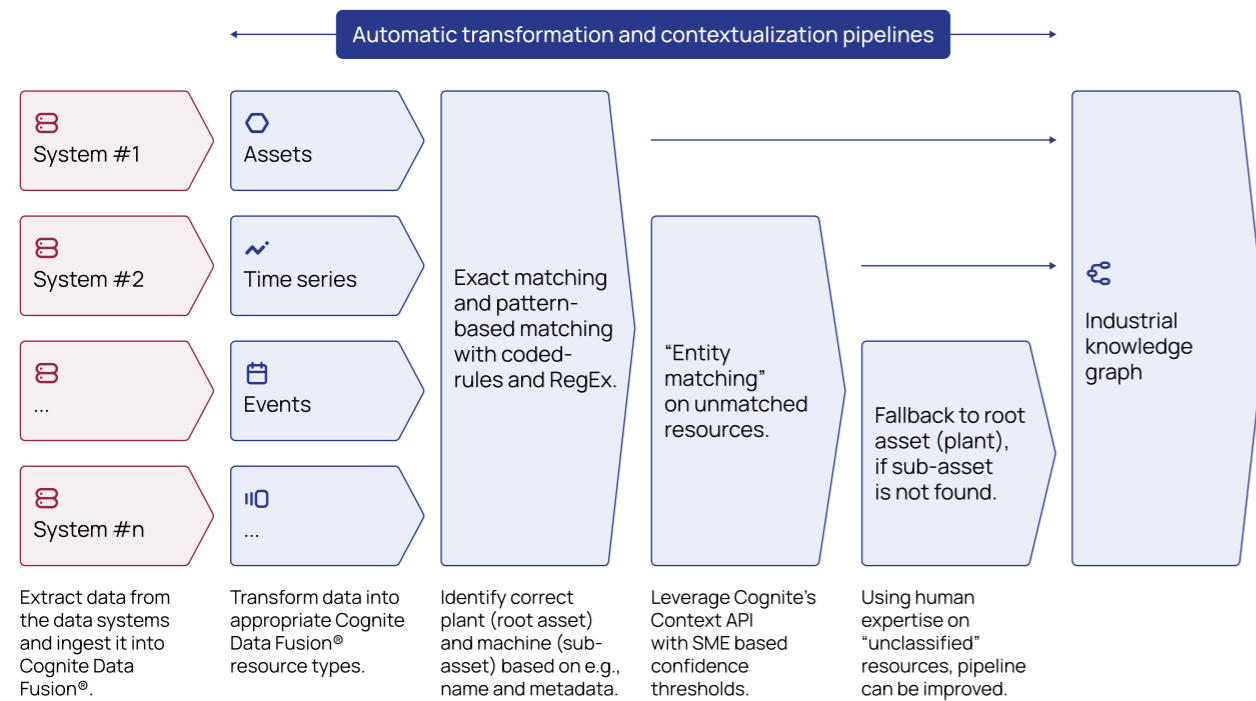
Entity matching is the swiss-army-knife of contextualization and matches string data properties, such as equipment names, descriptions, metadata, etc.

For example, you can match assets with time series to related work orders to their particular nodes in a 3D model. Matching signals must be present for different data sets to be successfully matched and appended into the reference and application data models. The entity matching model uses AI to find matches when there are similarities between the strings and does not return suggested matches for unrelated entities. Even with weaker matching

signals, a data contextualization engine provides enormous value by organizing, structuring, and governing a Subject Matter Expert's (SMEs) intensive data contextualization work. For example, spreadsheets and CSV files are no longer needed to map relationships manually. Moreover, generative AI further enhances the accuracy of these models by improving and ensuring the accuracy of entity matching models to better understand the semantic relationships between the previously siloed sources.



Example of how “entity matching” is leveraged in a contextualization pipeline to automate the formation of the industrial knowledge graph



Create Interactive Diagrams

Contextualization capabilities extend beyond entities and can also build interactive engineering diagrams/P&IDs from static PDF source files. Contextualization of diagrams and drawings discovers and validates relationships from these previously flat PDF files automatically without writing any code.

This contextualization is achieved by identifying tags within the PDFs using OCR and smart algorithms to find the correct tags linking, e.g., to files or equipment in those files. The detected instances are then stored as annotations alongside the file and link to the relevant data, available to navigate from the PDF to the relevant asset or process data.

Contextualize Imagery Data

Whether captured with cameras, drones, or a mobile phone, images and video data contain valuable information about the state of a facility over time (inspections, maintenance routines, etc.). However, often utilizing this data remains a challenge. Establishing contextualization pipelines to extract relevant information from these data types requires several industry-relevant and ready-to-use services. For example, computer vision can identify people, safety equipment, spills, reading analog gauges, and much more. These services should exist both as APIs and SDKs and can be used in automated pipelines for analyzing imagery data with the power of AI.

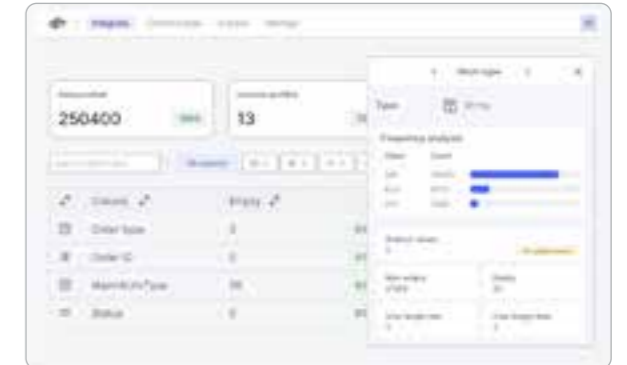
Create Document Classifiers

We can enrich documents such as standard operating procedures, inspection rounds, or OEM manuals with more context through connections to data sources such as Sharepoint and local file directories. This contextualization will transform static pages into dynamic, interactive documents with live links to assets or processes identified in a document. Plus, contextualization with generative AI enables natural language document search, enabling users to quickly find the correct information without manually scanning documents.

The ability to automate the creation of contextualization pipelines with AI-based algorithms across all of these different data types shortens the data contextualization process from months to days. Using AI-based algorithms to build an industrial knowledge graph efficiently avoids six-plus months of upfront effort and increases time to value. Additionally, with live contextualization pipelines, the effort to scale is significantly reduced, and data can be applied to solve many use cases across many sites without increasing the size of teams to manage this foundation dramatically.

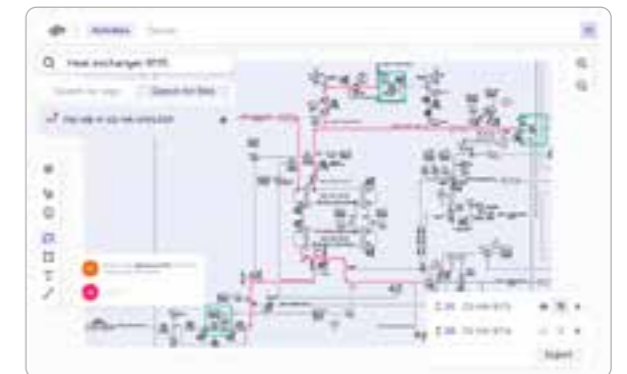
Contextualize events to assets

E.g. connect shop orders, work orders, and alarms, to correct site, prod. line and equipment



Contextualized documentation

E.g. connect tags in P&IDs to associated assets, time series, work orders, and documents etc.



Contextualized 3D models

E.g. filter for and visualize where work orders are located on plant

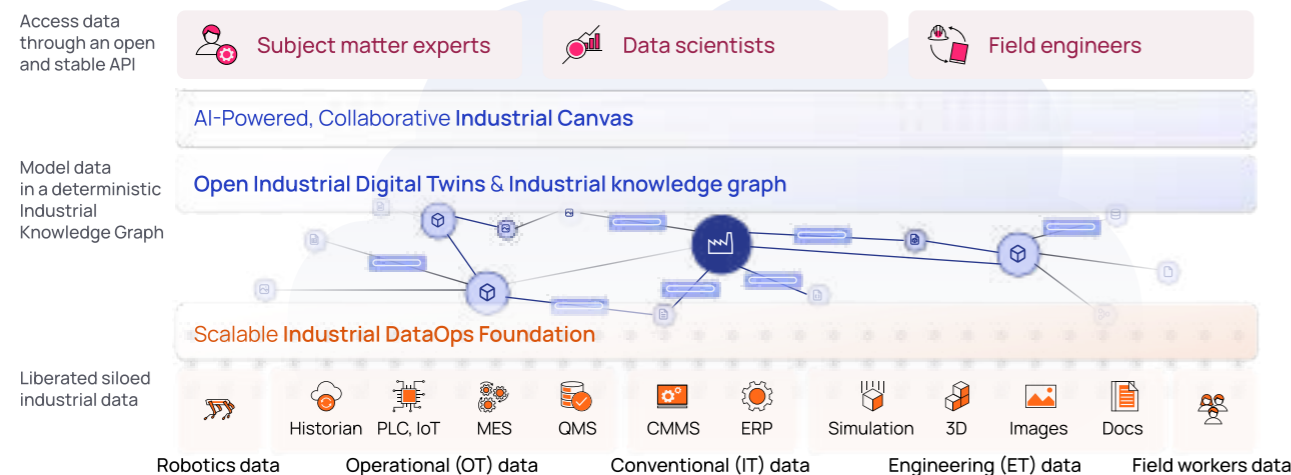


Contextualization, Data Models, and Digital Twins

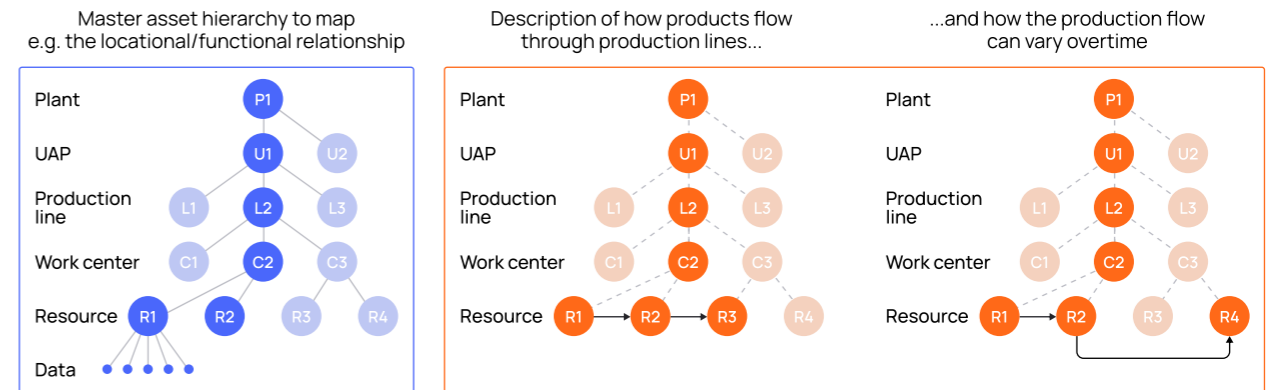


Contextualizing Data Into an Industrial Knowledge Graph

Contextualization is necessary when building an open, flexible, labeled industrial knowledge graph to represent your operations. Data modeling makes it easier for all stakeholders to find the necessary information and view and understand relationships between data objects.



The knowledge graph enables you to capture various data perspectives



Take, for example, someone developing a production optimization application across multiple production lines.

They need a robust domain API that provides instant access to a data model, which contains all the relevant data structured to accurately reflect the process while ensuring performant querying, regardless of where that data originates from or where it is stored now. And they need it all in domain language, not in the language of databases. So in this simplified example, the data model powers this domain API.

Simple aggregation of digitized industrial data is a significant step forward from the silos and inaccessibility that often plague large enterprises. However, to provide simple access to complex data, the variety of industrial data types must be accounted for, and the semantic relationships that drive scalable utilization of this data must be incorporated to support interactive user experiences.

Codifying this context as an industrial knowledge graph is vital to enabling consistent, deterministic navigation of these meaningful relationships.

With the industrial knowledge graph as the foundation, data is understood and structured to meet the specific needs of users or use cases. For example, asset resources commonly originate from a maintenance system, and the hierarchical asset structure of the maintenance system can define how the asset resources are organized.

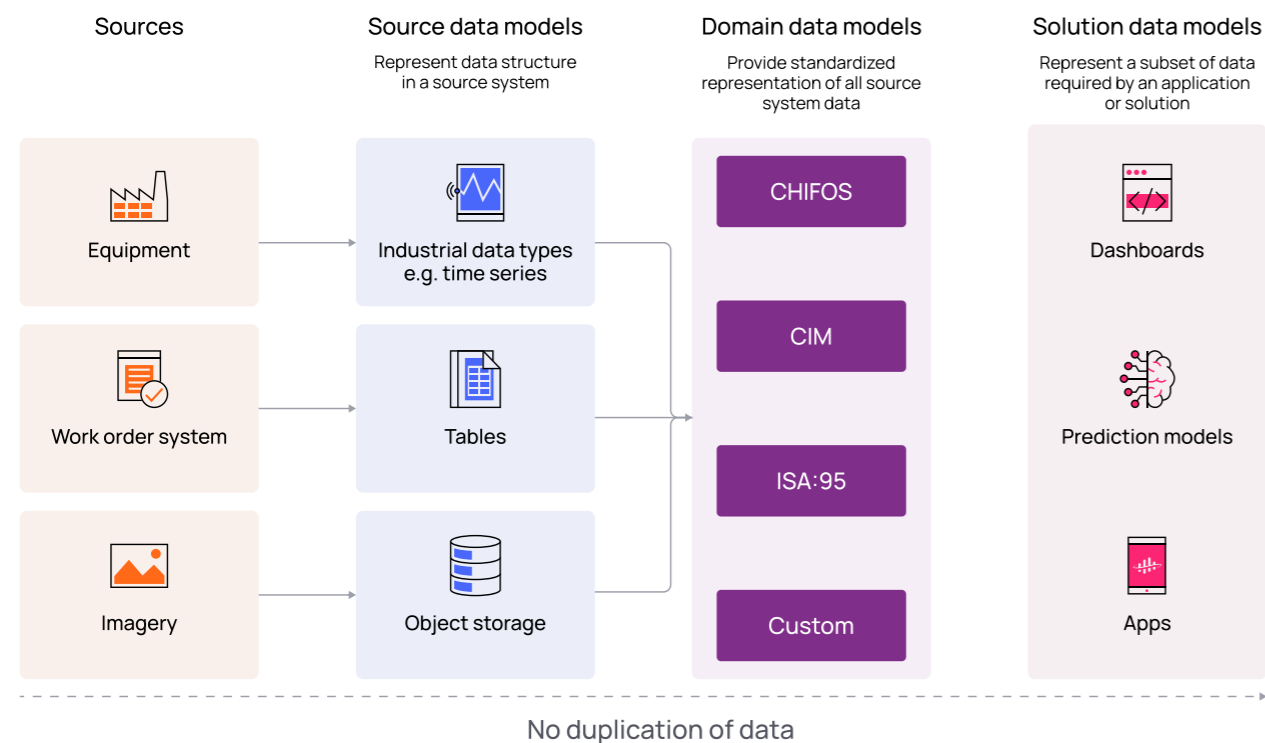
An asset hierarchy is ideal for addressing use cases related to asset performance management (APM). The relationships resource type also allows the organization of assets in other structures besides the standard hierarchical asset structure. For example, assets can be organized by their physical location, where the grouping nodes in the hierarchy are buildings and floors rather than systems and functions. Building a model with this structure opens the possibility to solve new use cases like product traceability, where the physical connections of the assets through the production process must be known.

Data becomes an asset, liberated from its silos, with reusable analytics and scalable models, shareable across many users. This industrial knowledge graph encourages data reuse by creating a user-friendly architecture. By leveraging data effectively and rapidly, the organization can address business opportunities quickly and at scale.

Not One to Rule Them All

To transform operations, industrial companies must build tens of data-driven solution and then scale them across hundreds of production facilities.

Scalability helps enterprises break the Proof of Concept (PoC) purgatory cycle.



While some companies have lighthouse sites, where technology allows different teams to work harmoniously, achieving this performance state commonly takes one to two years. If an organization has 50 locations in total, do they have 50-100 years to achieve the same level of performance at each of those 50 sites?

To address this lack of speed, the industry needs an approach that combines domain and industrial data expertise into a single product, enabling data reuse to develop many tailored solutions rapidly. Data modeling is a core component of turning siloed data into scalable solutions.

Physical, industrial systems are complex to represent, and no single representation will work in all the different ways to consume the data. The solution to this complexity is standardizing on a set of data models that contain some of the same data but allow you to tailor each model and add unique data.

For this, we need a data modeling framework that allows different perspectives of the same data to be clearly described and reused. Data models for industry can exist at three levels:

1. **Source Data Model** — data is liberated from source systems and made available in its original state.
2. **Domain Data Model** — siloed data is unified through contextualization and structured into industry standards (CFIHOS, ISA-95, etc.).
3. **Solution Data Model** — data from the source and domain models that support specific solutions.

The different layers enable value creation holistically and on a per-project basis. While the source data model liberates data from various source systems, it should also be queryable through the same API interfaces. The domain data model allows for a higher level of entropy and the representation of evolving ontologies, whereas the solution data model is much more rigid while at the same time allowing for true scalability across two dimensions:

- **The scalability of one solution** is ensured through an automated population of solution data model instances made possible by the contextualized relationships in your industrial knowledge graph, e.g., scaling a maintenance optimization solution across the entire asset portfolio.
- **Scalability across a portfolio of solutions** is enabled by immediate access to a wide range of data sources and the fact that application requirements for data are decoupled from the representation in the domain data model. This decoupling allows for use cases to be solved that require different levels of data granularity, e.g., plant-level maintenance optimization vs enterprise-level strategic planning.

As a result, enterprises can break free of the PoC purgatory cycle and focus on use case innovation that delivers in production at scale.

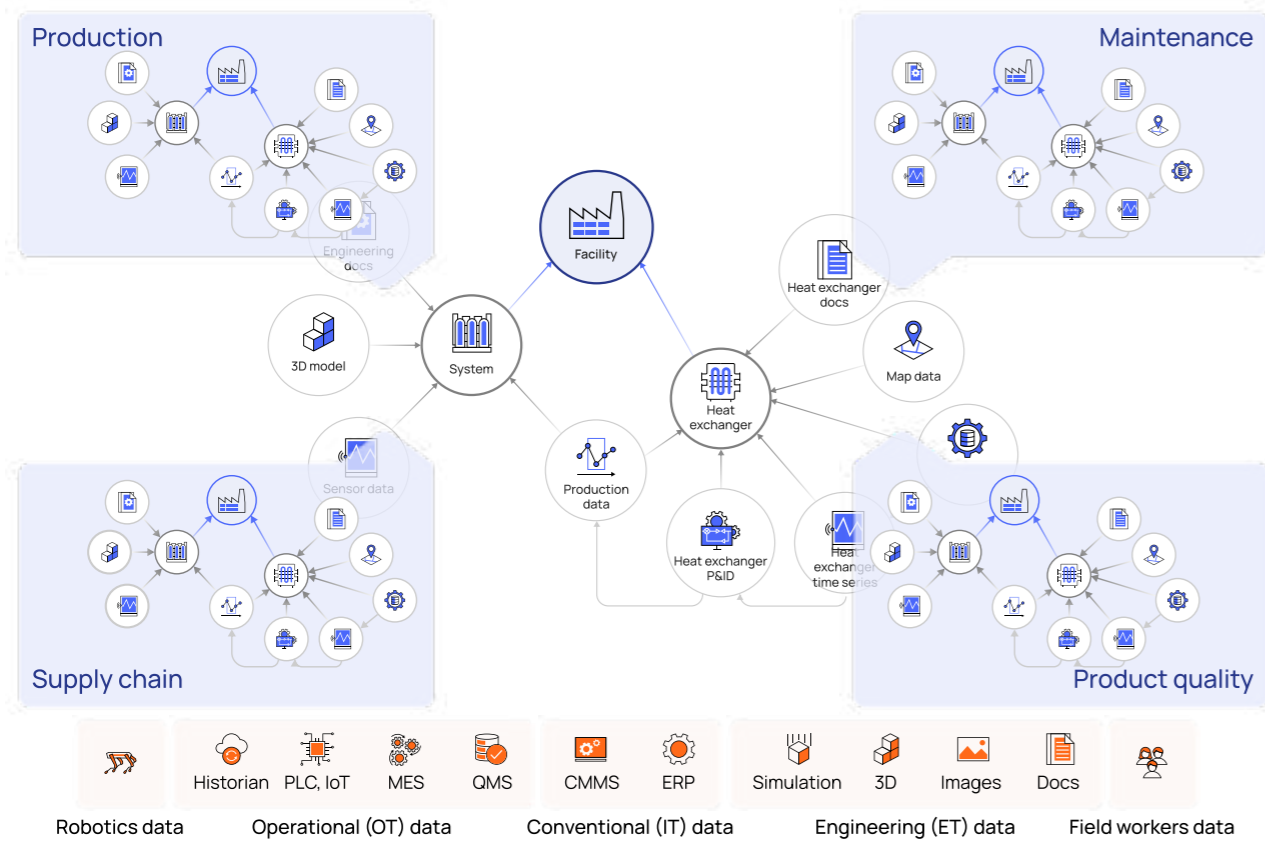
Data Models Work Together in a Digital Twin

The most prevalent application of data modeling is to unlock the potential of industrial digital twins. The advantage of data modeling for digital twins is to avoid the singular, monolithic digital twin expected to meet the needs of all and focus on creating smaller, tailored twins designed to meet the specific needs of different teams. The industrial knowledge graph acts as the foundation for the data model of each twin and provides the point of access for data discovery and application development.

Industrial companies can enhance the overall understanding of their operations by creating relationships across OT, IT, and engineering data using contextualization pipelines to develop an open industrial digital twin.

An industrial digital twin is the aggregation of all possible data types and data sets, both historical and real-time, directly or indirectly related to a given physical asset or set of assets in an easily accessible, unified location. The collected data must be trusted and contextualized, linked to the real world, and made consumable for various use cases.

Digital twins must serve data to align with the operational decision-making process. As a result, companies need not a single twin but multiple twins tailored to different decision types. For example, a digital twin for supply chain, one for various operating conditions, one for maintenance insights, one for visualization, one for simulation—and so on.



The above graphic shows that a digital twin isn't a monolith but an ecosystem. What is needed is not a single digital twin that perfectly encapsulates all aspects of the physical reality it mirrors but rather an evolving set of 'digital siblings' who share a lot of the same 'DNA' (data, tools, and practices) but are built for a specific purpose, can evolve on their own, and provide value in isolation.

To support that ecosystem, industrial companies need an efficient way of populating all the different kinds of digital twins with data in a scalable way.

Contextualized data unlocks valuable insights for operators, increasing understanding and improving operations processes. With contextualized data, industrial companies find it easier to examine their assets across multiple levels, from individual sensors to complex models. Equipped with virtual representations of real-world assets reflecting real-time data, operators can identify and mitigate problems that have been present but invisible for decades.

“Data management platforms help firms streamline the data capture, handling and contextualization processes, while accelerating time to value, cloud deployment, and scalability.”

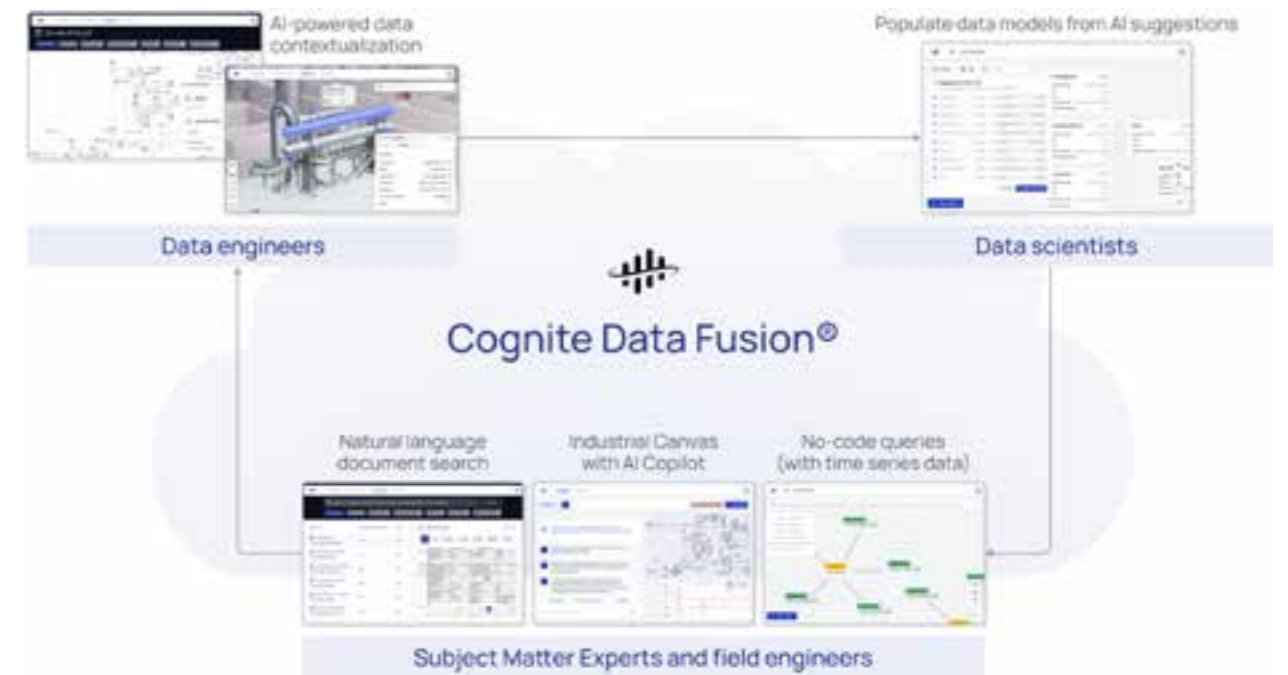
(Verdantix Green Quadrant Report)

Industrial Generative AI Requires Data Contextualization

Contextualized data shortens the time to business value and provides a pathway to scale in many industrial performance optimization applications and across advanced analytics workstreams. Additionally, access to contextualized data allows Subject Matter Experts (SMEs) to become more confident and independent when making

operational decisions or working on use cases with data scientists and data engineers. While contextualization helps solve the complexity of industrial data for users, we must not underestimate the impact of providing industrial context to generative AI solutions.

Our early exposure to LLMs through ChatGPT and Dall-E made the industry quickly realize the incredible generative capabilities of this latest advancement in deep learning AI.



Broadly, AI is accelerating contextualization, auto-populating data models, and simplifying the use of data for those who know how to code and those who don't. All of this results in accelerated time to value, improved operations, and an even easier way for SMEs, operators, and data scientists to consume and work with data.

The ability of AI to accelerate contextualization and to populate data models has been discussed throughout this book. The final piece of a more intuitive data experience is the benefits of generative AI.



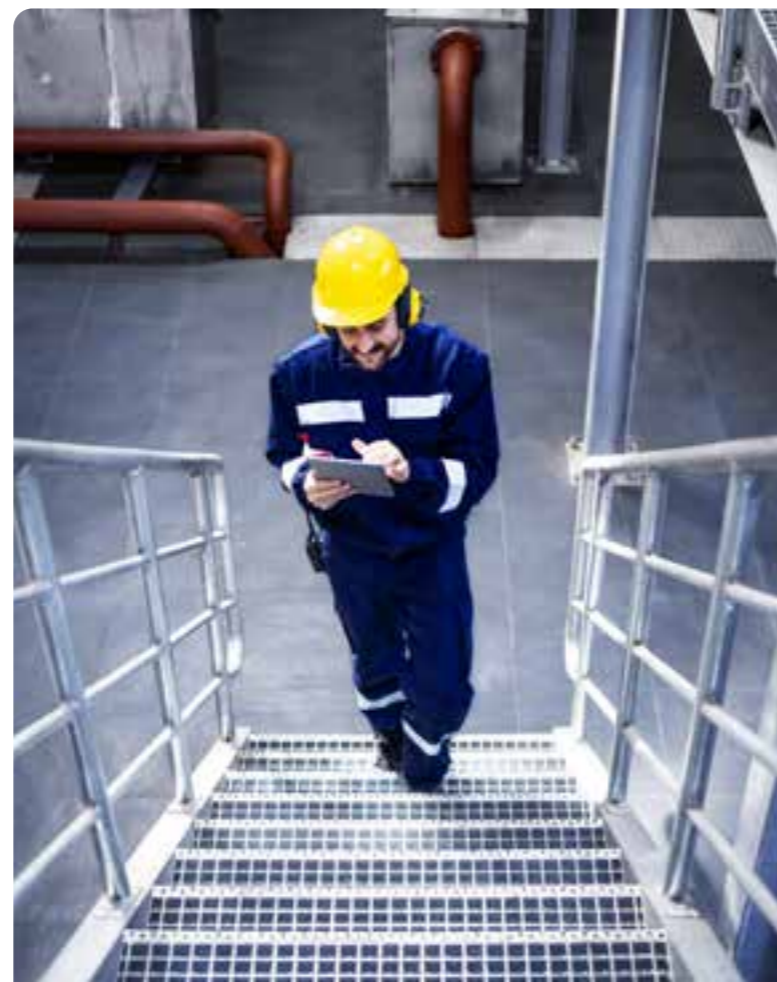
Q&A with a CTO



Jason Schern,
Field CTO at Cognite

He has spent the last 25 years working with some of the largest discrete manufacturing companies to dramatically improve their data operations and machine learning analytics capabilities.

For the past year, much of his time has been spent talking directly to operators or manufacturers about how generative AI can and should be incorporated into their digital transformation strategies. Here, we cover some of the most common questions he gets from operations and IT teams in the field.



Q: What are some specific risks companies aren't yet considering when incorporating generative AI models that haven't undergone extensive testing?

Jason: There are a couple of critical aspects of large language models (LLMs) that are important to recognize:

- They are trained on a large corpus of text, including web content, books, software source code, etc. By its very nature, this content includes context, which significantly impacts an LLM's ability to make predictions when 'reading' new content or generating content in response to prompts.
- LLMs are probabilistic and will return responses consistent with both the original training and reinforcement learning used to build the model and any inputs provided for consideration when responding to a prompt.

As a result, it is not the level of testing that drives the risk profile, but the amount/type of context.

For industrial data and industrial use cases, the risk profile is determined by the level with which a general-purpose LLM (i.e., Open AI, Bard, etc.) can return a deterministic response based on highly contextualized information. When a general-purpose LLM has access to highly contextualized information, the 'reasoning engine' can significantly reduce the probability of hallucinogenic responses to provide both deterministic and verifiable responses.

Q: Many businesses feel pressure to board the generative AI train. What are some key considerations companies should weigh up before investing? Which companies/industries are better placed to invest in developing their own generative AI products, and which companies/industries should outsource?

Jason: This is highly dependent on the goals and profile of the business. For businesses in creative pursuits or content generation, jump on the generative AI train with confidence and a high sense of urgency.

Generative AI is designed and optimized for content generation and creative outputs, especially as manifested in some of the most recent LLMs. These businesses can outsource their investments (i.e., fully leverage publicly available LLMs).

However, for many other businesses and industries, accuracy and validation of responses might be critical to ensuring operational efficiency and safety. These entities will need to invest in generative AI pursuits that can better leverage the high-level reasoning of generative AI to provide trustworthy, deterministic responses.

This need for trust typically leads to approaches where contextualized industry-specific data can be provided as inputs alongside prompts to generally available LLMs—or specialized LLMs can be fine-tuned utilizing highly contextualized, industry-specific data sets for additional training.

Because these approaches depend on proprietary industrial data, these companies will be unlikely to outsource information that can become a competitive advantage.

Q: Generative AI's adoption via ChatGPT has been unprecedented—where do we go from here? Are we headed in the right direction? Which industries possess the most significant potential for harnessing benefits from generative AI?

Jason: We are still very early in the adoption curve of generative AI. ChatGPT is an excellent showcase for the capabilities of generative AI to the broadest possible audience with a non-existent learning curve due to the natural language approach to generating prompts. We will begin to see compounding effects where generative AI and other technology curves will lead to exponential advancements in most fields. The list of possible directions is too extensive to address here, but the most significant impacts will go far beyond chatbots. It's the application of generative AI with other technologies that, once combined, will yield the greatest innovations.

Some obvious examples can be experienced today in areas such as autonomous driving or flight (Tesla/SpaceX), where multi-modal generative AI joins photogrammetry, telemetry, and other technologies to outperform humans in performance and safety. Other examples will come from medicine and materials science, where new treatments and new materials will be discovered.

In industry, we'll see tremendous advancements in operational efficiency and safety as we combine generative AI with robotics, photogrammetry, and other optimizations that previously were limited to the rare subject matter expertise of highly experienced professionals.

Yes, we are headed in the right direction. However, the speed and breadth of innovation are so unprecedented in technology adoption cycles that it will be challenging to follow these advancements and prepare for the implications.

It is difficult at this early stage to identify where the greatest potential lies.

However, companies that can organize and contextualize proprietary data will significantly benefit from the innovations and efficiencies generative AI will enable.

The eternal adage of 'garbage in, garbage out' not only applies to generative AI, but it is also the greatest determinant of the extent to which a company can successfully leverage the integration of generative AI capabilities into their business.

Q: Numerous conversations are happening around how generative AI should be regulated. Do you think about how it should be handled and which areas pose the greatest risk? What considerations should organizations bear in mind regarding the potential impact of government regulations and standards on their future utilization of generative AI?

Jason: I may be a bit of a contrarian on this point. The 'cat is out of the bag,' and any attempts at regulation will be unenforceable in a world where these models are already so ubiquitous across the globe.

Any regulation attempt would either need to be globally enforceable (impossible) or run the risk of placing limits on investments in specific countries.

The implications of regulatory limitations would risk slowing or limiting advancements that will shift to other countries. The national security implications of falling behind in these advancements will make it highly unlikely that any meaningful regulations will be passed. However, I do feel it likely that major players (Open AI, Microsoft, Google, etc.) will begin to seek legislation that provides some 'regulatory capture' to protect their positions.



Q: In your opinion, what is the biggest misconception about generative AI?

Jason: The two main misconceptions—at least at this early stage—are:

1. Experiences based on ChatGPT can feel like LLMs understand text. They do not. LLMs expose statistical probabilities resulting from the data context and fine-tuning used during model training. This dependency on context necessitates such a large corpus of data for generic model training.

When this is understood, we realize that LLM performance in providing non-hallucinogenic responses is greatly impacted by the quality of the data and the level of context inherent in that data.

The better the data (context-rich), the less is required for model training (significantly lower costs for model development or fine-tuning), and the greater the chances for generative AI to provide more deterministic, verifiable responses.

2. The second misconception is that generative AI can understand complex industrial data just like text, books, web content, programming languages, etc. This is also not true. There are several critical differences between industrial data and text-based content, but the most significant difference is the lack of context. Contextual cues primarily influence the statistically driven models of generative AI. These contextual cues are inherent in the content itself. Industrial data lacks these contextual cues in the data. Sensor

data from machinery provides constant information through a timestamp and sensor value. There is no context inherent in this data. The machine this sensor belongs to, the work orders that have previously been performed, operating conditions, operating throughput, maintenance history, and other critical contextual information are not included in the sensor data.

All of this context is critical and either needs to be utilized in training a specialized generative AI model or included as input when using a general-purpose AI model. Context and data quality matter more than ever.



Q: What are the most significant headwinds and tailwinds you anticipate will affect generative AI over the next six to 12 months?

Specifically in industry, the availability of high-quality, contextualized data will be the most significant barrier to utilizing generative AI's ever-growing capabilities and innovations. Those with it will out-pace their rivals in what generative AI can provide. Tesla is an obvious example of this that we can see now. Their access to high-quality, contextualized data from their customers' cars will allow for advancements and innovations that other automakers are unlikely to replicate.

The cost of compute (availability of GPUs and cost/prompt) will be a limiting factor in many use cases. This challenge is temporary, and I expect market dynamics will resolve this in six to 12 months, but we will not be able to ignore these throttling effects early on.

And lastly, I suspect there will be many failed attempts to leverage generative AI in various use cases. Many of these endeavors will suffer from not understanding what generative AI is doing and what it needs to succeed (high-quality contextualized data). We will see many failed attempts that have at their core a level of wishful thinking not grounded in technology.

To avoid this pitfall, start by solving your data problem before jumping straight into generative AI initiatives (and note; these can be done in parallel!).



AI is the Driving Force for Industrial Transformation



From Digital Maturity to Industrial Transformation

There are two discomfoting truths within digital transformation across asset-heavy industries:

1. Digitalization proof-of-concepts (PoCs) are commonplace. Real Return on Investment (ROI) isn't.
2. Organizations invest billions in cloud data warehouses and data lakes. Most data ends there, unused by anyone for anything.

Value capture reality

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



ABUNDANT digitalization ambition and necessity



MOST industrial companies have run at least a handful of AI proofs-of-concept (POCs)



MANY POCs are considered successful



FEW POCs ever make it into production



VERY FEW POCs end up providing substantive ROI (by becoming proofs-of-scale)

While nearly every organization is talking about digital transformation, the use of data, scaling, and time to value, only some in the industrial world are actually reaping the benefits.

There is no shortage of data in any industrial company, but there is a general lack of understanding on how to extract it, bring it together, and use it in a valuable way. At the heart of this data-driven value dilemma lies a confluence of challenges, ranging from the technical (“How can we best organize our diverse and fluid data universe?”) to the operational (“How can we create new information products and services?”), to the financial (“How can we treat data as an asset?”), to the human (“How can we improve data literacy and ensure digital solution adoption in the field?”).

Though more and more industrial operations data is readily available in the cloud, Chief Data Officers (CDO) are confronted with the hard reality that moving data to the cloud is not even a third of the journey to value. As Forrester (2021) put it elegantly, “Data has no value unless the business trusts it and uses it.”

Data and analytics leaders need to capitalize on the value of data fully. However, many aren't sure how. So what is required to fix the industrial data problem? Industry has put much effort into new digital tools and capabilities to create free information flows from their vast amounts of industrial data. But the data systems and mechanics are not even connected enough to approximate Google-like search for front-line staff.

The convergence of data and analytics has made Industrial DataOps an operational necessity.



DataOps focuses on delivering business-ready, trusted, actionable, high-quality data available to all data consumers.

- **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, infused with AI, promises to improve the time to value, quality, predictability, and scale of the operational data analytics life cycle. It provides the opportunity to offer data science liberation within any product experience while simultaneously allowing subject matter experts (SMEs) to acquire, liberate, and codify domain knowledge through an easily accessible and user-friendly interface. It's a stepping stone to a new way of managing data in the broader organization, enabling it to cope with growing data diversity and serve a growing population of data users.

Industrial DataOps and AI for Industry

While generative AI can help make your data 'speak human,' it doesn't speak the language of your industrial data on its own. As discussed in previous chapters, a strong industrial data foundation is required to remove the risk of hallucinations. Generative AI can improve data usage, but trustworthiness requires a robust data foundation with a powerful contextualization engine.



As Verdantix recently put it, “DataOps-focused vendors are uniquely placed to harness the power of AI/ML technologies such as LLMs for asset-heavy industries—taking advantage of industry-specific data ingest, transform, governance and contextualization tools to provide a unique grounding plane for hallucination-prone LLMs.”⁵

Industrial DataOps platforms help data workers deploy automated workflows to extract, ingest, and integrate data from industrial data sources, including legacy operations equipment and technology. It 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 data is then made available through specific application services for humans, machines, and systems to leverage.

Industrial facilities generate more data than ever with greater volume, velocity, variety, and visibility. Organizations are changing their approach to data and operations to meet digitalization priorities.

Gartner expects that by 2024, the use of synthetic data created with generative AI will halve the volume of real data needed for machine learning and that by 2026, code written by developers/humans will be reduced by 50% due to generative AI code generation models.

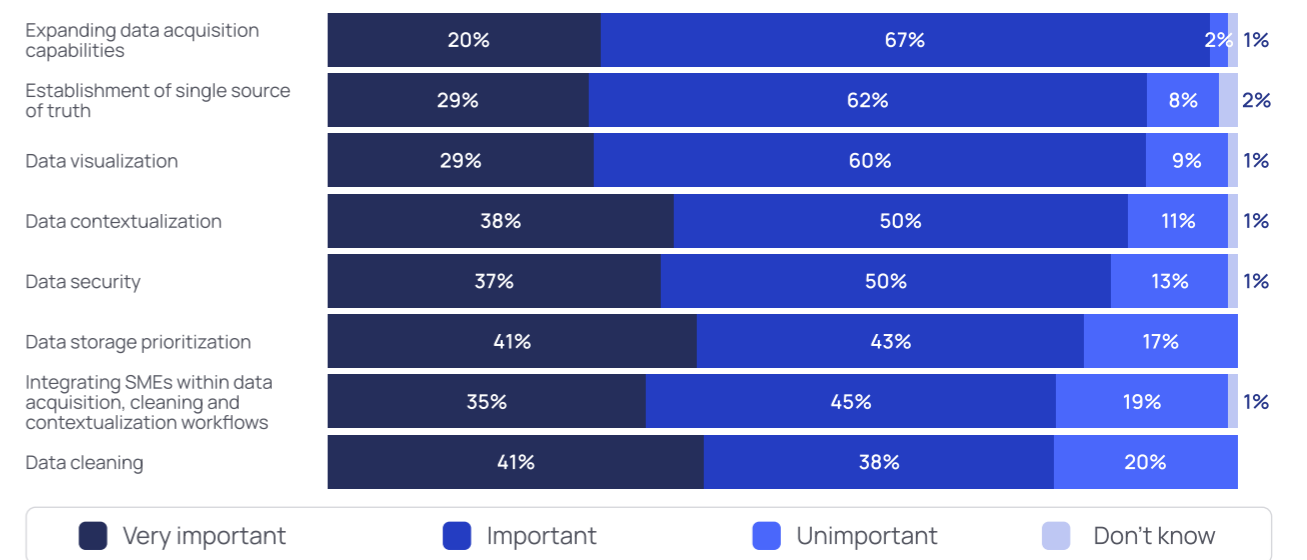
Digital mavericks have already noted what is in their best interest and are aligning their strategic technology purchases with their enterprise transformation strategy.

⁵ www.verdantix.com/insights/blogs/cognite-s-new-interface-gives-industry-its-first-real-glimpse-of-generative-ai-s-value →

Rapidly maturing generative AI technologies are poised to change how humans process complex information or situations. While SMEs and operators are getting used to new data infrastructure that helps them 'Google' complex industrial data, AI will further simplify the human-to-data interface.

Industrial DataOps and Data Management Priorities for 2023

How important are the following data management activities for your firm over the next 12 months?



AI Can Deliver Untapped Value for Asset-Heavy Enterprises

To adequately extract the value of industrial data insights, it's essential to make operationalizing data core to your business strategy. Data must be available, useful, and valuable in the industrial context. Operational teams need a robust data foundation with a strong data context and interpretability backbone, all while applying generative AI to accelerate workflows that optimize production and make operations more efficient.

Efficient Data Management and Improved Data Accessibility

A strong data foundation is required to remove the risk of 'hallucinations' and increase AI 'readiness.' An Industrial DataOps foundation 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. Using AI to enable rapid ingestion and contextualization of large amounts of data 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.

Augmented Workflows and Process Improvements, Driving Innovation at Scale

Using generative AI-powered semantic search, what used to take your process engineers, maintenance workers, and data scientists hours of precious time will take only a few seconds. With the guidance of generative AI copilots, users can generate summaries of documents and diagrams, perform no-code calculations on time series data, conduct a root cause analysis of equipment, and more. Time spent gathering and

understanding data goes from hours in traditional tools to seconds. Now users can spend more time driving high-quality business decisions across production optimization, maintenance, safety, and sustainability.

Rapid Development of Use Cases and Application Enablement

Too often, digital operation initiatives get trapped in 'PoC purgatory,' where scaling pilots takes too long or is too expensive. Use an AI-infused Industrial DataOps platform to shorten the time to value from data by making PoCs quicker and cheaper to design and offering operationalizing and scaling tools. These copilot-based approaches leverage the power of natural language to understand and write code based on published API documentation and examples to support development processes.



Generative AI further improves ML training sets of ML models by generating synthetic data, enhancing the data set used for training, enhancing process efficiency, and optimizing production. Some common use cases in asset-heavy 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.

Enterprise Data Governance as a By-Product and Personalized AI tools

By having a strong Industrial DataOps foundation, you can then empower users to adapt AI models to cater to their specific requirements and tasks, using generative AI to enhance data onboarding, complete with lineage, quality assurance, and governance, while a unique generative AI architecture enables deterministic responses from a native copilot.

On top of that, Industrial Canvas overcomes the challenges of other single pane-of-glass solutions, which often over-promise capabilities and are too rigid with prescribed workflows. This prevents users from working with the data how they choose, by delivering the ultimate no-code experience within a free-form workspace to derive cross-data-source insights and drive high-quality production optimization, maintenance, safety, and sustainability decisions.

If implemented successfully, an AI-augmented data platform provides consistency and ROI in technology, processes, and organizational

structures, with better operations data quality, integration and accessibility, and stewardship. It should also enhance data security, privacy, and compliance with tracking, auditing, masking, and sanitation tools.





Democratizing Data: Why AI-Infused DataOps Matters to Each Data Stakeholder

Extracting maximum value from data relies on applying advanced models to produce insights that inform optimal decision-making, empowering operators to take action confidently. Turning insight into action is, in a nutshell, what we mean by operationalizing data into production for value.

But for every one person that can 'speak code,' there are hundreds who cannot.

Generative AI will change how data consumers interact with data. It 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. It is about removing the coding and scripting and bringing the data consumption experience to the human user level.

Making the data speak human is the only way to address the Achilles heel of practically all data and analytics solutions, especially those for heavy-asset industry verticals. These organizations face many challenges; an aging workforce, extreme data type and source system complexity, and very low classical data literacy among SMEs; those needing data to inform their daily production optimization, maintenance, safety, and sustainability decisions.

Why it Matters to Executives

Priorities	Challenges	Value
Financial performance and profits	ROI on investment, cost of downtime	Enable data-driven decision-making that allows focus on the highest ROI activity at any given time
Innovation and solution delivery	The measures and KPIs used in decision-making are entrenched in showing short-term value	Serve as a bridge to increased digital maturity, as 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
Reduce inefficient processes that waste time and effort	Scaling on assets and equipment; projects are slow to deploy and done in isolation	Align and bring together formerly isolated subject matter experts (SMEs), cultures, platforms, and data deployed by IT and OT teams to improve operational performance through unified goals and KPIs
Impact reputation, sustainability and future viability	Poor reputation and sustainability	Empower organizations to operate with greater precision, and track the correct metrics to reduce the impact on the environment
Having the best workforce	Scaling the number of people creating solutions; lack of skill set	Enable existing non-professional data users to perform some data management tasks, and drive value for the enterprise, while preserving their valuable accumulated knowledge and experience
Remaining relevant and competitive	Missing the digital transformation revolution	Leverage data effectively and rapidly to answer questions and gain some insulation from market volatility
Enabling digital transformation	Confusion about how to deliver digital transformation and what it means for both management and workers	Enable the organization to meet the need for fast-moving innovation by providing consistency and ROI in technology, processes, and organizational structures, with better operation data quality, integration and accessibility, and stewardship
Strategic change in culture and vision	Reactive culture is an obstacle to growth	Recognize data as an enterprise asset and build a path towards digital maturity through tools and processes so that digital ways of working become effortless across a broad range of stakeholders

Why it Matters to IT and Digital Teams

Priorities	Challenges	Value
Data preparation and integration	Legacy systems. Complex integrations and dependencies between multiple data sources	Minimize existing data silos and helps simplify the architecture to support rapid development and deployment of new analytics
Create solutions for operations by turning insights into actionable advice	There's too much data, with no context; challenge in making data usable and getting models in operation	Offer 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
Automate workflows	Scale across insights, solutions, sites	Facilitate industrial equipment and processes data models and templates—it talks domain language and scales models from one to many
Minimize the hurdles in cross-functional collaboration	Multiple hand-offs that are error prone and increase risks	Provide a unified approach to streamline end-to-end workflows from data preparation to modeling and insights sharing
Driving implementation of Industry 4.0	Difficult to explain business-wide benefits to senior decision makers	Connecting data users with disparate operational data sources helps bridge those divides on the path to use-case operationalization

Why it Matters to Domain Experts

Priorities	Challenges	Value
Optimization of process around quality, throughput, and yield	Lack of insights or tools to make quick and correct decisions around maintenance and production	Leverage domain knowledge and human expertise to provide context and enrich data-driven insights, and further develop machine learning models that utilize data to improve planning processes and workflows
Inefficient processes waste time and effort	Work in isolation and without full possession of all the data and the facts	Provide the capabilities domain experts need to support self-service discovery and data orchestration from multiple sources
Resource management	Reluctancy and slow adoption of new tools and tech in old ways of working	Empower business functions to use data and supports the digital worker
Improve planning process and workflow across assets and equipment	Manual tasks are time-consuming and prone to errors	Industrial DataOps infused with AI enables your site to solve today's challenges with a pathway towards more autonomous systems and sustainable growth

Solving the industrial data problem is critical in realizing value from industrial digitalization efforts. Benefits can be measured from streamlined APM workflows, improved SME productivity, optimized maintenance programs, and real-time data efficiencies, such as:

- **Productivity savings due to improved SME efficiency.** DataOps provides data accessibility and visibility, transforming the way data scientists and SMEs are collaborating.
- **Reduced shutdown time.** The opportunity cost of large, industrial assets being out of production is significant. Using a digital twin and better component data visibility, SMEs are able to safely minimize shutdown periods when data anomalies arise.
- **Real-time data access enables improvement in productivity.** Live data access enhances operational flexibility and decision making by increasing site safety, improving predictive maintenance, and raising machine performance.
- **Optimized planned maintenance.** Cognite Data Fusion® creates contextualized data to optimize planned maintenance by analyzing and interpreting available resources, workflows, and component life cycles.
- **Energy efficiency savings.** Intelligent data can be used to reduce energy use and therefore operational costs.
- **Optimization of heavy machinery and industrial processes.**
- **Health and safety.** Reduce the amount of human movement through potentially dangerous 'hot' areas, reducing risks to employee health and safety.
- **Environmental, social, and governance (ESG) reporting.**



Use Cases for Industrial Generative AI



Copilots:

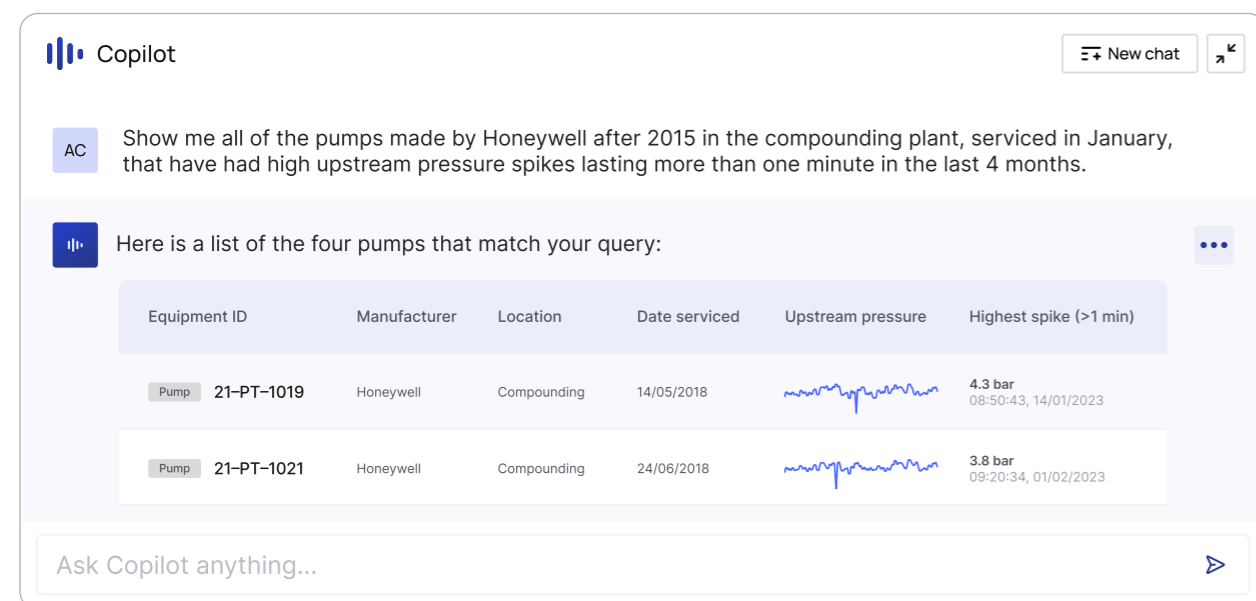
Enhancing Self-Service and Humanless Workflows

There are two discomfoting truths within digital transformation across asset-heavy industries:

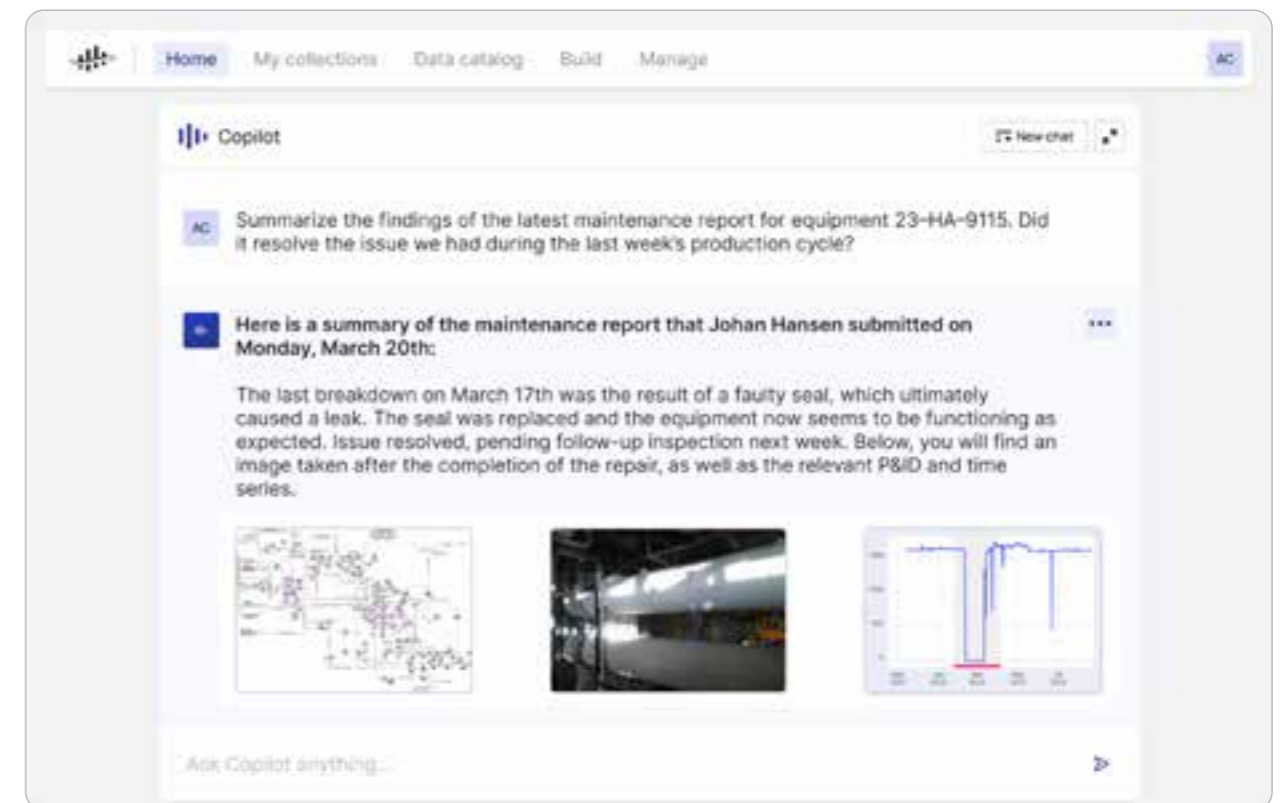
Whereas a process engineer might spend several hours performing 'human contextualization' (at an hourly rate of \$140 or more) manually—again and again—contextualized industrial knowledge

graphs provide the trusted data relationships that enable generative AI to accurately navigate and interpret data for operators without requiring data engineering or coding competencies.

With generative AI-powered semantic search, what used to take process engineers, maintenance workers, and data scientists hours of precious time now takes only a few seconds.



Take, for example, a copilot approach. Because LLMs such as ChatGPT understand and can generate sophisticated code in multiple languages (i.e., Python, JavaScript, etc.), users can prompt the LLM with a site-specific question. The LLM can then interpret the question, write the relevant code using an open API, and execute that code to return a response to the user (see the below image for example).



These copilot-based approaches leverage the power of natural language to understand and write code based on published API documentation and examples. This level of automation is impossible with data lakes or data warehouses where, without a contextualized industrial knowledge graph, there are no API libraries that can be used as a reliable mechanism to access rich industrial data.

Industrial data can also be provided, in context, directly to the LLM API libraries available from OpenAI, LangChain, and others, allowing users to

leverage the power of the LLM's natural language processing in conjunction with proprietary data. This database can include numerical representations (embeddings) of specific asset data, including time series, work orders, simulation results, P&ID diagrams, and the relationships defined by the digital twin knowledge graph. Using these open APIs, a prompt can be sent to the LLM along with access to a proprietary embeddings database so that the LLM will formulate its response based on the relevant content extracted from your proprietary knowledge graphs.

Industrial Canvas: Interact with Industrial Data in a Modern, Open Workspace

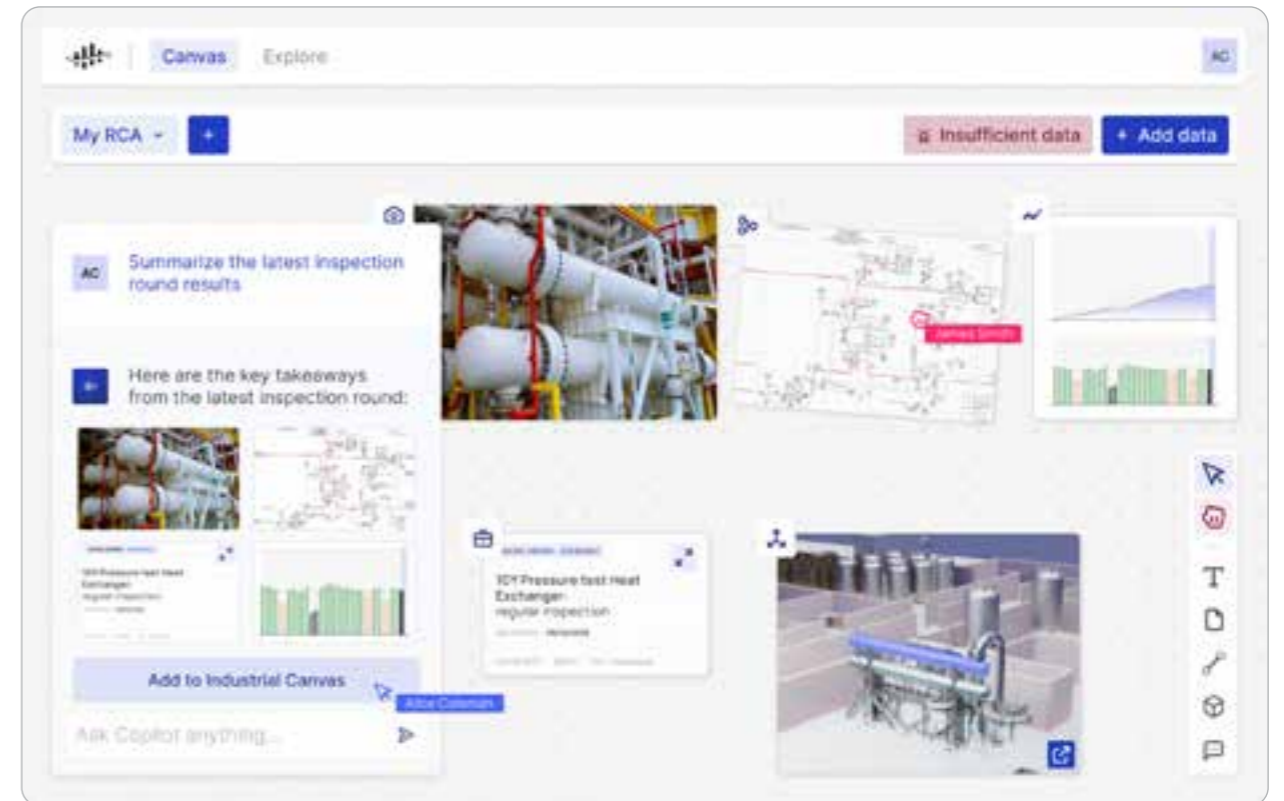
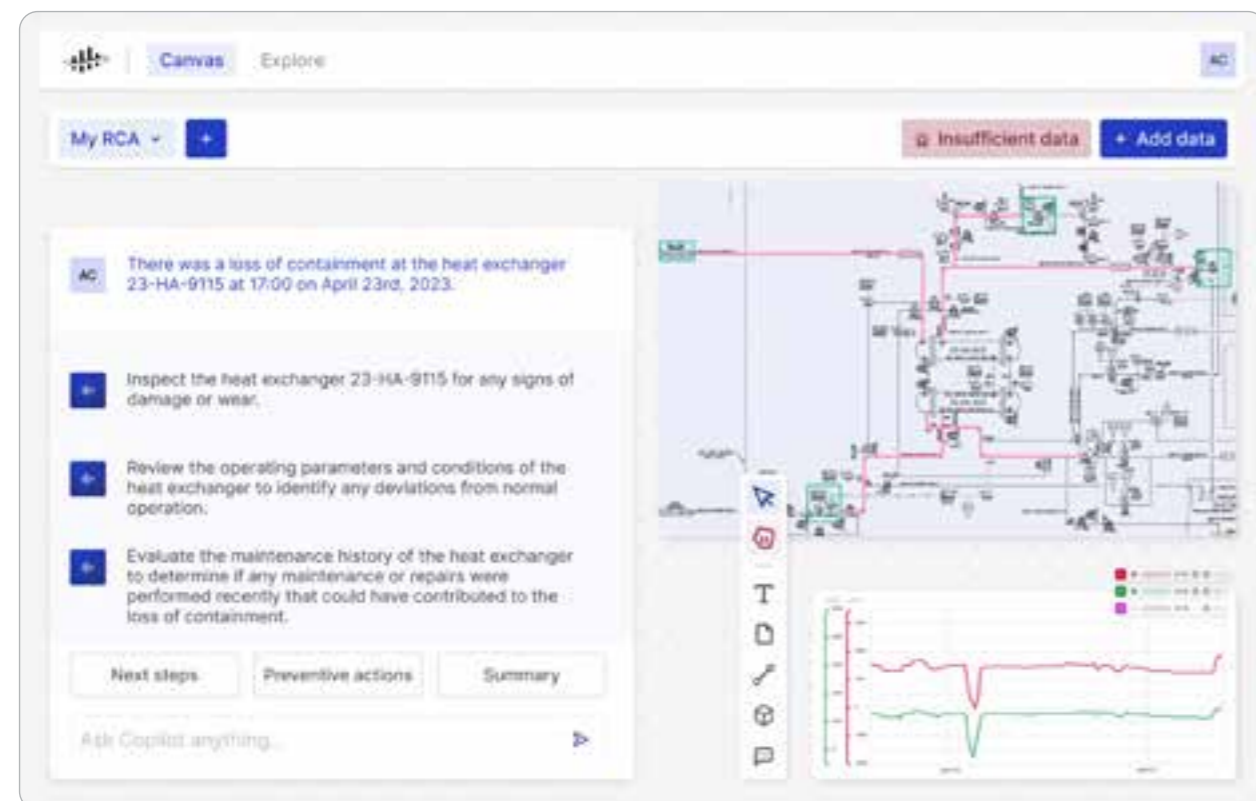
For every one person that can 'speak code,' there are hundreds who cannot.

Users deserve a new way to interact with industrial data, one that delivers an ultimate no-code experience within a free-form workspace to derive cross-data-source insights and drive high-quality production optimization, maintenance, safety, and sustainability decisions.

With an intuitive, user-centric tool revolutionizing data exploration and visualization, SMEs and

operators should be able to make cross-data source insights available without relying on data scientists, data engineers, and software engineers to build tailored case solutions. As a result, this enables users at every level of the organization to spend less time searching for and aggregating data and more time making high-quality business decisions.

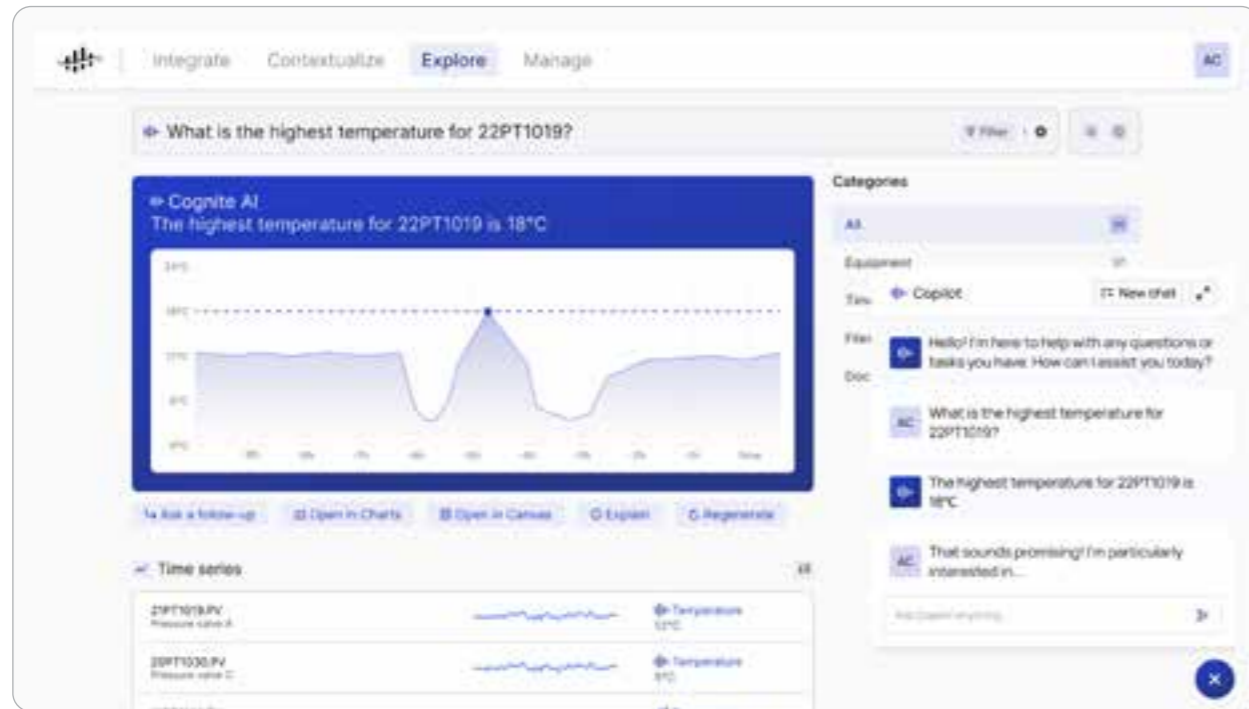
With an industrial knowledge graph making data available in an open workspace, users can view all data types in one place, including documents, engineering diagrams, sensor data, images, 3D models, and more—data that matters to industrial end users, including process engineers, maintenance planners, reliability engineers, machine operators, technicians, and others.



Users can generate documents and diagram summaries, perform no-code calculations on time series data, and conduct root cause analysis of equipment with the guidance of the copilot.

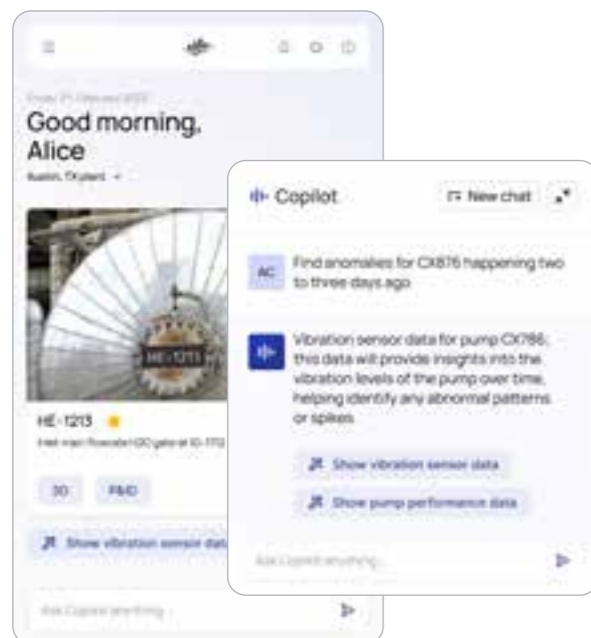
SMEs deserve to work with their industrial data in a collaborative environment where they can share workspaces, tag other users, and share insights as comments. This approach overcomes the single pane of glass solution, which often over-promises as a framework. Plus, this approach is often difficult to collaborate with and too rigid, preventing users from working with the data how they choose.

This concept, captured as a product, is a perfect example of how contextualized data generates immediate business value and significant time-savings in many industrial performance optimization applications and across advanced analytics workstreams. Access to contextualized data allows Subject Matter Experts (SMEs) to become more confident and independent when making operational decisions or working on use cases with data scientists and engineers.



Real-World Examples from Cognite

Let's look at other examples of AI applications for industry.



Gen AI-Powered Data Exploration

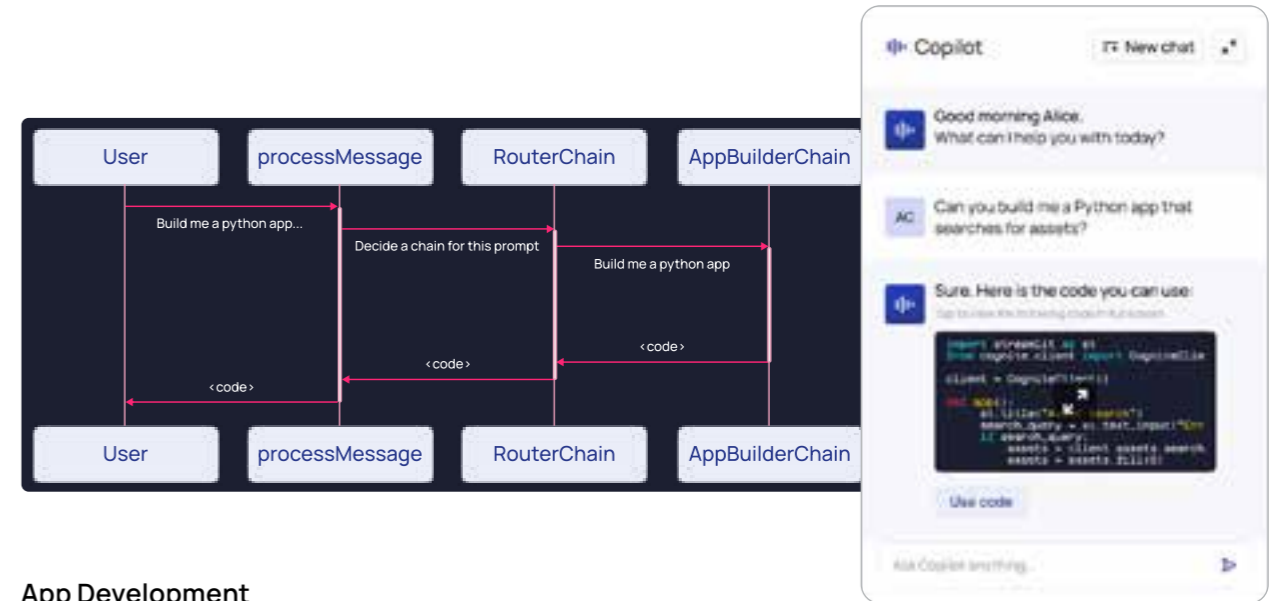
Problem: Domain experts don't have a deep understanding of their data models and how to query across data types ('relational queries').

Solution: A natural language as an alternative to traditional search, wherein the expert can interact with a copilot to refine a search and get clear feedback to find the data they need for day-to-day tasks quickly.

Document Insights

Problem: Field workers need to understand equipment specifications buried in huge documents.

Solution: A copilot that summarizes the document's relevant paragraph(s), along with links to document locations, to accelerate decisions.

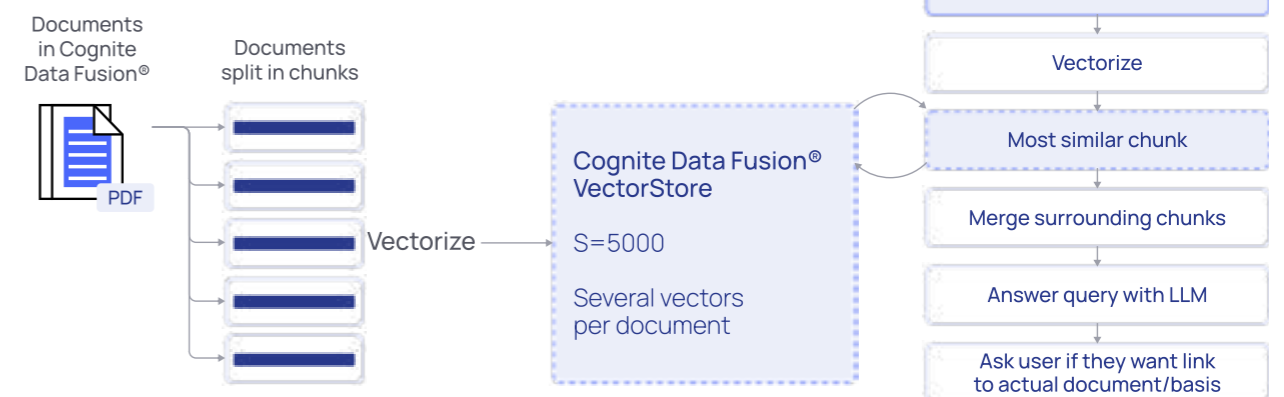


App Development

Problem: Developing industrial applications is time-consuming and costly. It is challenging to get the right data streams that provide the right insights to end users in a scalable way.

Solution: Prompt engineering, or using NLP to discover inputs that yield the most desirable results, to replace many software applications and bypass the traditional application development cycles.

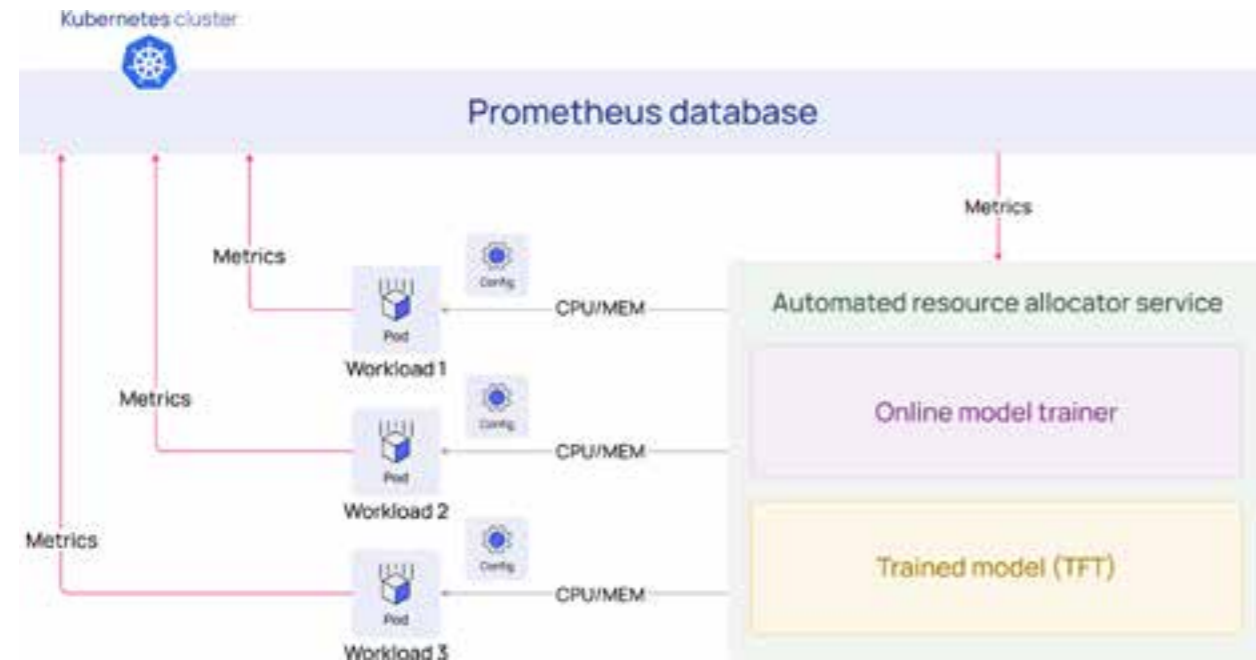
Solution description



ISO Compliance

Problem: Industrial companies must comply with numerous standards to ensure regulatory compliance and manage operational risks, requiring significant resources and coordination of siloed knowledge.

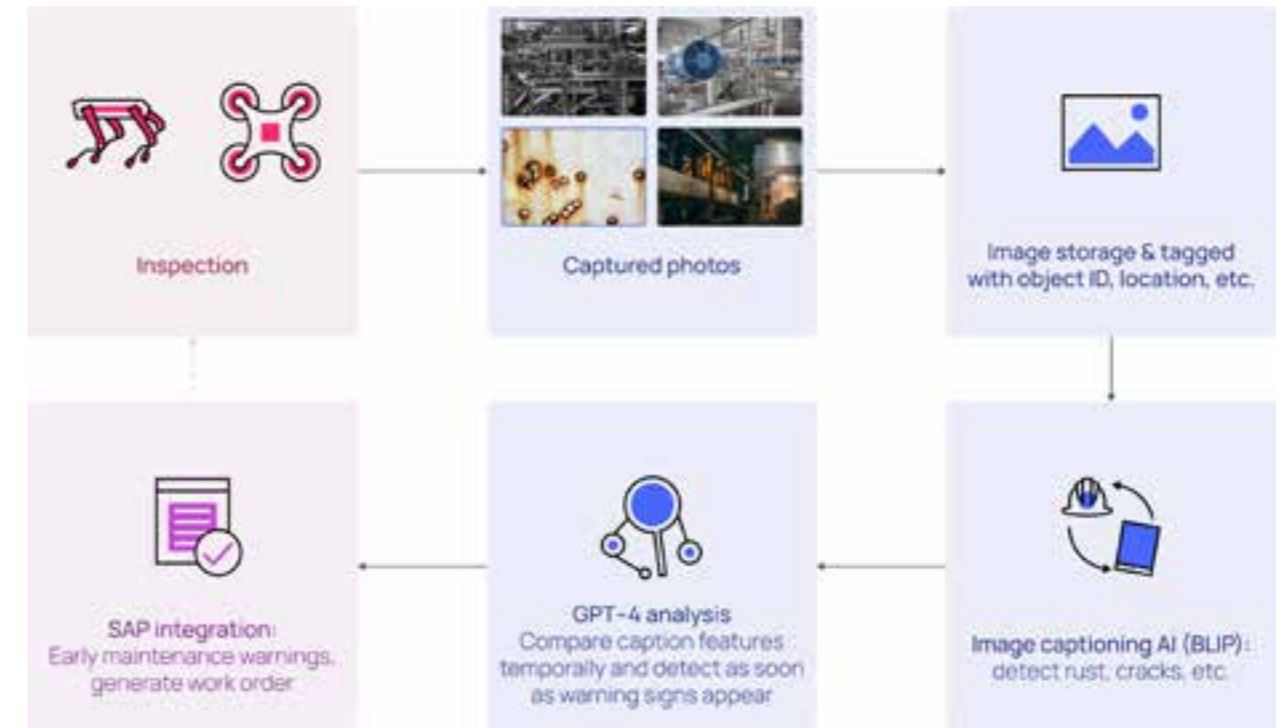
Solution: Copilot for simplifying the compliance process for industrial companies through GenAI and vectorization.



Resource Utilization

Problem: Efficient resource utilization is crucial for businesses relying on Kubernetes. Traditional resource prediction methods do not capture the dynamic nature of Kubernetes workloads, resulting in suboptimal resource allocation and increased cost.

Solution: A state-of-the-art AI tool for dynamic resource allocation that uses previous resource utilization (memory, CPU, etc.) data to forecast resource utilization in the immediate future. The solution can learn from the workloads and make intelligent scaling decisions and actions to allocate resources dynamically for k8s workloads.



Rust Detection

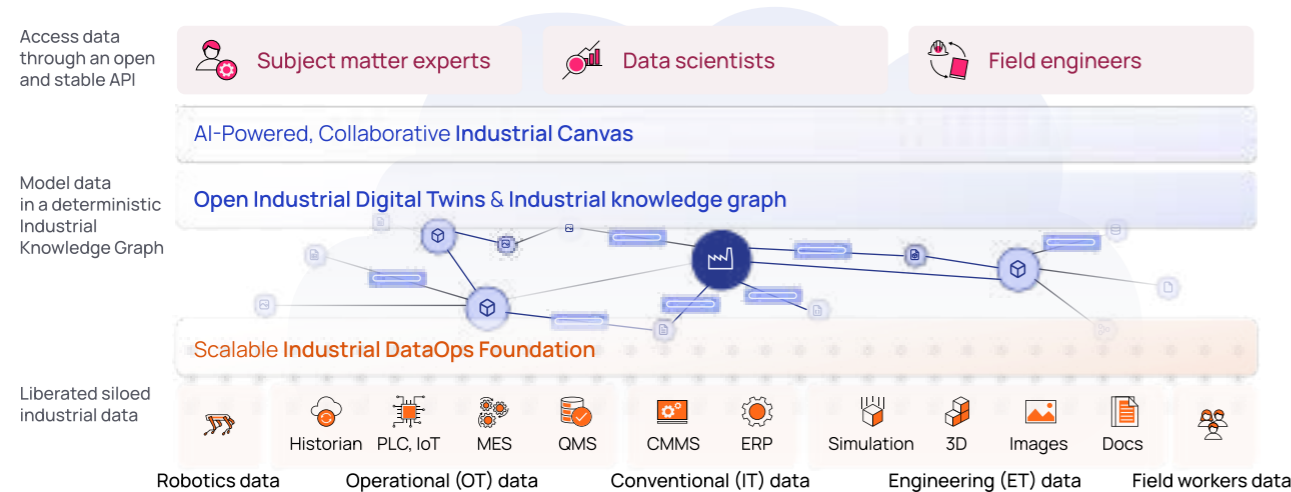
Problem: Rust can corrode and degrade equipment, machinery, and pipelines, leading to reduced functionality, decreased performance, and, in severe cases, complete failure.

Solution: Generative AI-powered Autonomous Operations incorporating robotic inspection rounds with computer vision and GPT-4 analysis for early rust detection and automated work order generation.

In summary, every user can benefit from contextualized industrial data—from a rich, intuitive data exploration for SMEs, production managers, business analysts, and engineers to a 'data as code' experience for data scientists who prefer SDK experiences. Creating an industrial knowledge graph by contextualizing data is a new way of serving all data consumers—data and analytics professionals and SMEs, business, and engineering professionals—with the same 'real-time contextualized data at your fingertips' experience.

With the continued loss of domain-specific knowledge caused by the aging workforce, various data types and source system complexity, and painful user experiences amongst SMEs, data contextualization is the only way to address the Achilles heel of all data and analytics solutions.

Data contextualization empowers everyone with improved decision-making to enable operational improvements, immediate business value, and significant time-savings in many industrial performance optimization applications, asset performance management, root cause analysis, and advanced analytics workstreams. Data contextualization is equally vital for scaling use cases, which is critical to transforming operations.

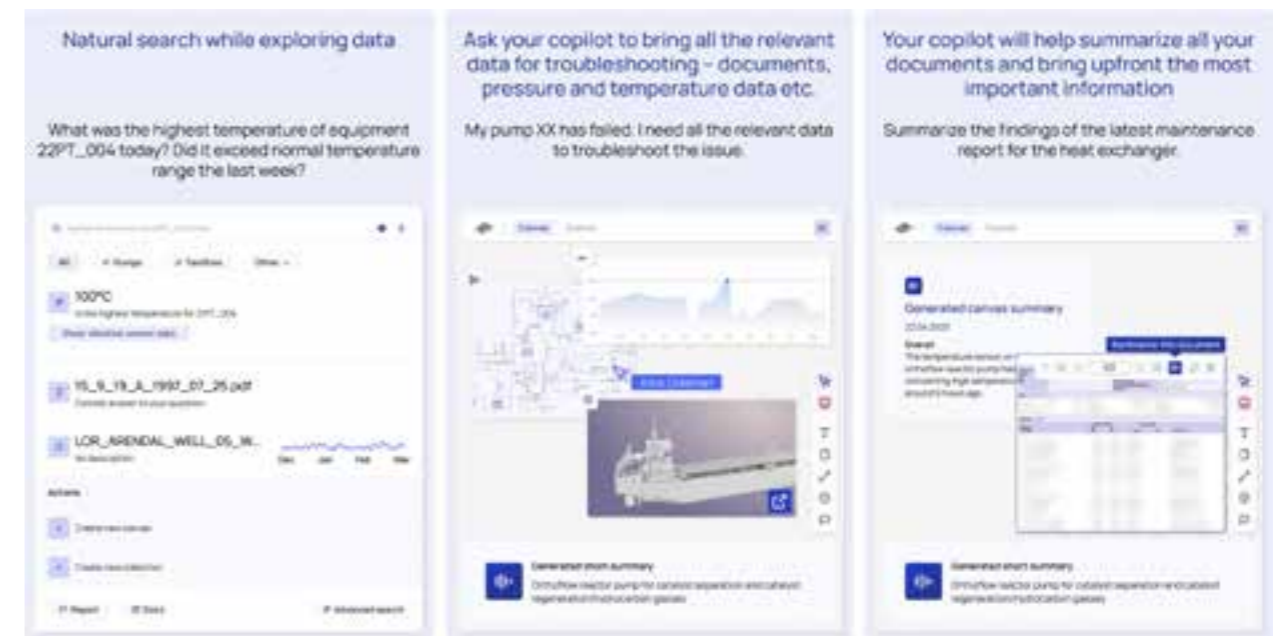


Simple access to complex industrial data will be achieved through an ultimate no-code experience within a free-form workspace to derive cross-data-source insights—a collaborative environment to share workspaces, tag other users, and share insights as comments. The goal is to provide an intuitive, user-centric tool that revolutionizes data

exploration, visualization, and analysis for process engineers, maintenance planners, reliability engineers, machine operators, technicians, and others, to build specific use case workflows without relying on data scientists, data engineers, and software engineers.

With the guidance of a generative AI copilot, users can also generate summaries of documents and diagrams, perform no-code calculations on time series data, conduct a root cause analysis of equipment, and more. Time spent gathering and

understanding data goes from hours in traditional tools to seconds. Now users can spend more time driving high-quality business decisions across production optimization, maintenance, safety, and sustainability.



Creating this experience is predicated on the ability to contextualize data rapidly at scale with AI-powered services that eliminate tedious manual contextualization and maintenance. Contextualization enables the industrial knowledge graph that serves as the connecting fabric between data modeling, digital twins, and all open components. Combining an industrial knowledge graph with a unique generative AI architecture, all users have simple access to complex industrial data. In their language. On their terms.

In the next two chapters, we look at how contextualization, industrial knowledge graphs, and generative AI are driving innovation in robotics, autonomous industry, and asset performance management.

The Impact of Gen AI on Robotics and Autonomous Industry



Robotics in industry is nothing new. Early industrial robots were typically programmed using pre-defined instructions and could not adapt or learn from their surroundings. In recent years, intelligent robots—capable of performing complex tasks and adapting to dynamic environments—have become much more common.

These robots are equipped with advanced sensors, artificial intelligence algorithms, and machine learning capabilities. They can perceive their environment, interpret data, and make intelligent decisions to perform complex tasks with minimal human intervention, expanding their application in industry.

For example, it is relatively common at this point to see a robot like Boston Dynamics' Spot navigating through a manufacturing plant's premises, following predetermined or dynamically adjusted routes to conduct inspection rounds. It can record images, videos, and audio snippets of critical assets and process the collected data in real time to identify signs of corrosion, leaks, abnormal temperature readings, or other indicators of equipment malfunction or safety concerns.

More advanced deployments combine robotics with digital twin technology. Autonomous inspection rounds deliver precise measurements of geometric attributes, identification of structural changes, and visual documentation of an asset's condition, which is the foundation for building and updating digital twins.

The continuous data collection by robots allows for real-time monitoring and detection of deviations or anomalies, reducing unplanned downtime, optimizing maintenance schedules, and facilitating data-driven decision-making.

However, even with these advancements, truly autonomous operations face certain limitations that can impede its widespread adoption and effectiveness:

- **Lack of adaptability:** Autonomous systems are typically designed to operate within specific parameters or pre-defined scenarios—like adapting to novel environments or dynamic operational requirements—and deviations from these conditions can pose challenges.
- **Limited decision-making in complex Scenarios:** Traditional rule-based or pre-programmed approaches may not encompass the intricacies of complex decision-making processes, limiting the autonomous system's ability to respond appropriately in dynamic and uncertain environments.
- **Lack of contextual understanding:** Autonomous systems may need help understanding the broader context or interpreting ambiguous information, particularly when it comes to tasks that deal with unstructured data.

Generative AI has emerged as a crucial component in achieving fully autonomous operations. It directly addresses the above by enabling adaptability, enhancing decision-making capabilities, and improving contextual understanding:

- **Adaptability through learning and generation:** Generative AI learns from vast amounts of data, including diverse scenarios and edge cases. It can generate new data, simulate different situations, and refine its capabilities through continuous learning. This enables autonomous systems to adapt and respond to evolving conditions, enhancing flexibility and robustness.



- **Enhanced decision-making:** Generative AI analyzes complex data patterns and generates insights based on learned patterns. It can consider a wide range of factors, assess risks, and generate optimal responses in complex scenarios, enabling more intelligent decision-making.
- **Improved contextual understanding:** Generative AI's ability to learn from human collaboration, especially when combined with RAG technology, enables autonomous systems to interpret and respond to human instructions or queries more accurately. This enhances the contextual understanding of autonomous systems, making them more capable of handling diverse and non-routine situations.



By leveraging the power of generative models, autonomous systems can become more resilient, versatile, and capable of operating effectively in complex and dynamic environments. This paves the way for the broader deployment of autonomous technologies across various industries, driving innovation, efficiency, and safety.

For example, generative AI, coupled with other AI techniques such as computer vision and natural language processing, enables robots to perceive and interpret their surroundings, communicate with humans, and make context-aware decisions.

Furthermore, generative models enable robots to simulate and predict outcomes in complex environments. By generating synthetic data and running simulations, robots can explore different scenarios, refine their decision-making algorithms, and improve their performance in real-world tasks.

This iterative learning process brings robots closer to achieving higher autonomy and operational efficiency.

In a simple yet poignant example, Cognite's Co-Founder and Chief Solutions Officer, Stein Danielsen, developed a ChatGPT-powered copilot solution built on top of Cognite Data Fusion® to monitor the temperature of a coffee machine in our Oslo office.

The solution alerts Stein to any inconsistencies to the temperature. He can then use voice-recognition to quite literally talk to the solution to check temperature trends, review a schematic of the plumbing connected to the coffee machine, and inquire about different parameters to troubleshoot the problem. He can even ask the copilot to send a Spot robot to visually inspect the coffee machine, using its static camera and

computer vision to analyze the valve position, heat map of the piping, etc. Stein can do all of this remotely; allowing him to determine the most likely cause of the temperature fluctuation and alert maintenance to the most likely solution. Most notably, after multiple tests of this solution, the copilot learned from previous interactions with Stein and began to recommend the best process for troubleshooting the maintenance issue.

We can easily extrapolate the benefits of this type of solution beyond a simple cup of hot coffee. Generative AI allows subject matter experts to find and solve problems more effectively. When enriched with contextualized data from an industrial knowledge graph, generative AI can serve as a cognitive extension for robots to make them nearly independent, assisting humans in even more tasks, expediting decision making and root cause analysis processes, and significantly improving productivity and efficiency of operations.

Generative AI also facilitates the development of collaborative robotics, where humans and robots work together towards a common goal. By generating insights and recommendations, generative models can assist humans in decision-making, optimize task allocation between humans and robots, and enhance the overall efficiency of collaborative work environments.

It is also exciting to note how well AI robotic solutions scale. As soon as one robot learns something new, *all* other robots in the system immediately know it too. When we use domain experts to train the AI that teaches these robots, we have found a way to infinitely scale human knowledge. The robot will be the embodiment of the AI and the AI can also use the robot to see and understand the real world.

Generative AI is fast revolutionizing the robotics and autonomous industry, bringing us closer to achieving fully autonomous operations. And the potential benefits are immense. With generative AI as a catalyst, asset-heavy industries are poised to achieve higher levels of automation, productivity, and innovation, redefining how products are designed, produced, and delivered.



The Impact of Gen AI on Asset Performance Management



90% of Asset Performance Management is about Simple Access to Data

Asset Performance Management (APM) software is critical as industrial companies develop a safer, more efficient, and more sustainable industrial future. Companies have invested in APM solutions to address business essential drivers of value, namely:

- Extending asset life and maximizing labor productivity.
- Reducing cost and time of maintenance.
- Increasing uptime and minimizing unplanned events.

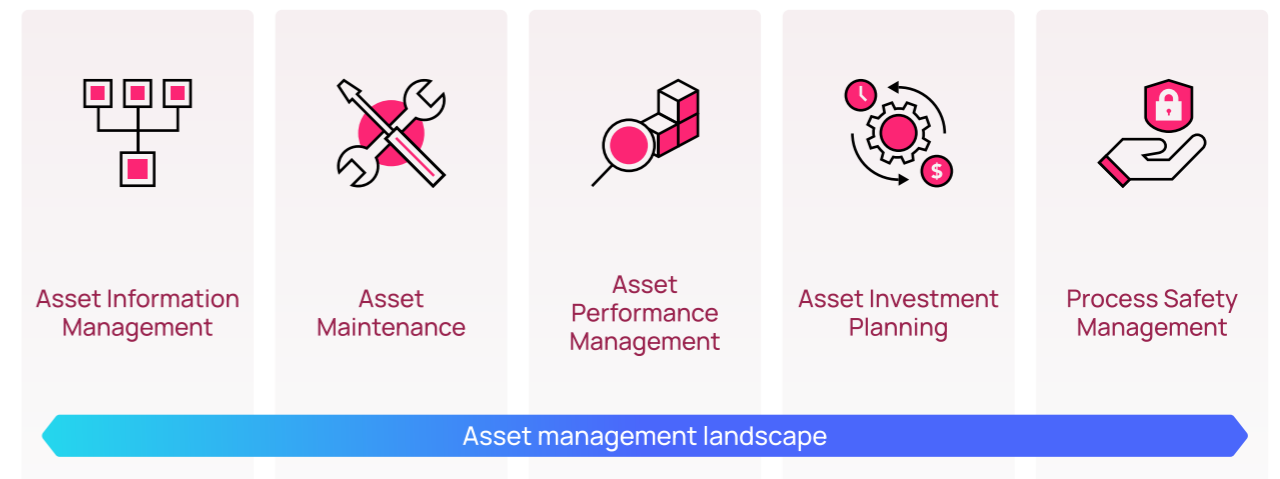
Historically, the solution for solving APM use cases has been to invest in siloed systems (e.g., ERP systems), niche solutions (e.g., IoT applications), or run multi-year 'lighthouse' projects with little-to-no ROI to show after months or years of data wrangling and deployment effort.

Other firms remain cemented in traditional paper-based or Excel-reliant processes, with data tucked away in filing cabinets or their digital cousins—data warehouses. Across the enterprise, collaboration is hindered, value is left unrealized, and organizations are never truly data-driven.



“Industrial software applications and associated asset content that monitor asset performance, predict failures, and synchronize with IT and OT systems to generate insights that help optimize production, reliability, maintenance, and environmental KPIs.”

Five market segments converge to form current asset management landscape



- 1. Lack of APM awareness & understanding** Strength: High

APM solutions and its benefits are still emerging in mature industries, whereas awareness remains limited in upcoming industries. Firms are yet to truly realize the true benefits of an APM solution.
- 2. Change management issues** Strength: Medium

Any technology implementation won't scale if either management or frontline workers are reluctant to change ingrained routines. With a lack of a skilled workforce and reluctance to change, adoption of APM has been hindered.
- 3. Complex architecture** Strength: Medium

Corporate leadership are prioritizing projects with quick ROI. Difficulties with integrating APM solutions with existing infrastructure is resulting in management to focus on other opportunities.
- 4. Data and cyber security threats** Strength: Medium

APM tools aggregate data from multiple sources and utilize cloud technology, presenting data management challenges and concerns w/r hacking and storing sensitive data.
- 5. Uncertain futures** Strength: Low

Concerns w/r to COVID-19, aging workforce, rising energy prices and political instability means firms are hesitant to invest.

The APM market has evolved from maintenance-related tasks to incorporate performance optimization, reliability, and environmental KPIs. (Verdantix, Green Quadrant) At the core of these issues is the industrial data problem: cross-department data silos, lacking context from undocumented data, inconsistent workflows, and a lack of successful cross-site scalability. Simply put, most APM strategies don't scale because people lack simple access to complex industrial data.

When people spend 80-90% of their time searching, preparing, and governing data, there's little time left to invest in gaining better, data-driven insights. APM initiatives will continue to fall short of their ROI expectations without a coordinated and collaborative data management strategy.

You may have succeeded with some strategies and technologies you have implemented, or even have a star site where your maintenance, reliability, and operations work in harmony. However, this probably took you one to two years to get to a specific performance state, and let's say you have 50 sites in total—do you have 50-100 years to achieve the same level of performance for each of those 50 sites?

If it takes six months to deploy one solution at one facility, how long will it take to deploy the same solution across all assets and all sites?

One of the main challenges of APM is to account for the messiness of real-world industrial data and the diverse range of use cases and users it must serve.



Asset Performance Management Intelligence: A Unified Approach to Achieve Continuous Improvement

APM achieves continuous improvement via establishing a continuous cycle of reliability, operations, and maintenance:

- Improving asset reliability and health monitoring: Surface and monitor functional reliability to inform risk-based decision-making.
- Optimizing activity planning and maintenance workflows: Ensure plan stability and maintenance process optimization.
- Performing efficiently executing work in the field and capturing new data: Efficient and safe field execution to complete the operational picture.

Asset-heavy organizations face numerous challenges when it comes to optimizing APM solutions, however, as it must serve different needs and use cases for a variety of asset-heavy organizations. Reliability engineers need solutions that enable them to accurately monitor equipment health and receive early warnings about potential failures to avoid breakdowns. For maintenance

workers, APM excellence means instant access to data so they can conduct maintenance on equipment based on its actual condition, as opposed to running its maintenance program on a set schedule. And your operations teams should be empowered to operate the 'digital plants of the future' with digital worker and robotics applications.

APM is a set of use cases to improve asset performance, ESG & sustainability, and optimize production

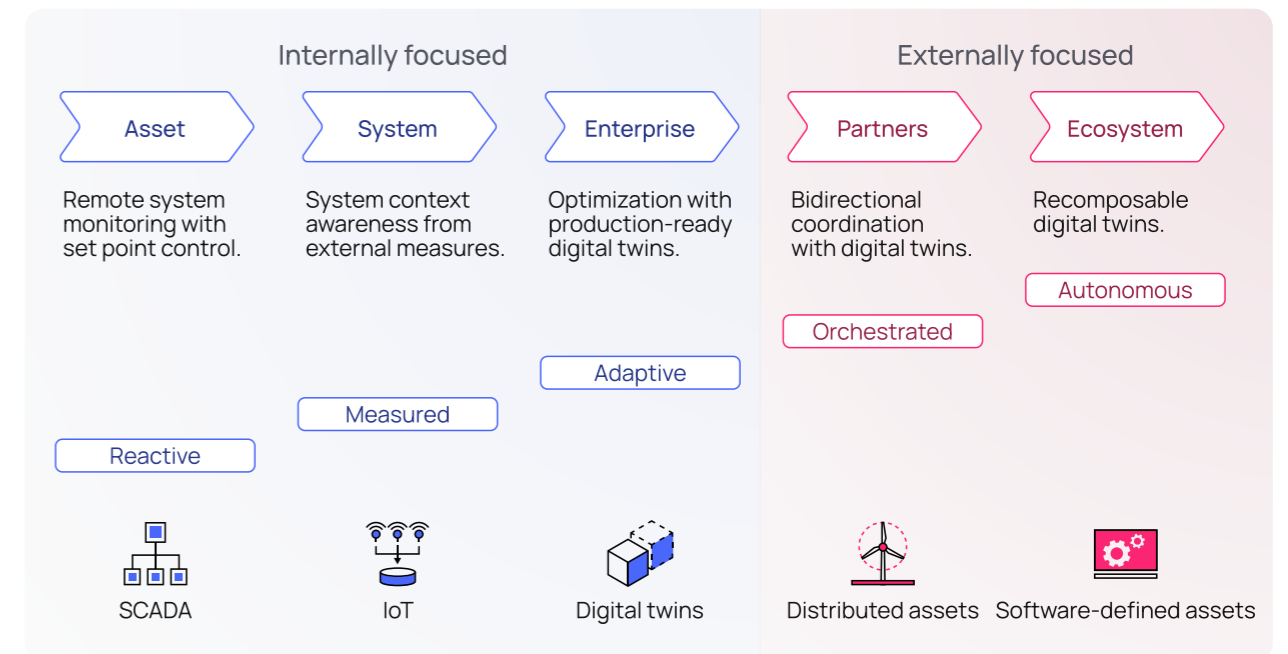


Source: Verdantix, Green Quadrant: Asset Performance Management Solutions 2022



APM intelligence is only achieved when the full value from these investments is scaled across an enterprise—not just with a single application or a one-off deployment. APM success depends on solving the industrial data problem, and that's why the best approach to APM is one that places an open industrial digital twin at its core.

A roadmap to adopt intelligent operations



Inspired by Gartner

An Open Industrial Twin for Asset Performance Management

To make sense of this data and use it to optimize asset performance, organizations need to invest in an open industrial digital twin that can capture the inherent entropy and ever-changing nature of physical equipment and production lines.

The open industrial digital twin provides the backbone to deliver the myriad of tailored, data-driven solutions that must be developed and scaled across users, production lines, and sites to harvest the value potential.

This digital twin also provides the necessary context for utilizing generative AI to provide deterministic responses to APM-related queries.

An open industrial digital twin is an aggregation of all possible data types and data sets, both historical and real-time, related to a given physical asset or set of assets in a single, unified location.

This digital twin provides a single, trusted data foundation for operations and analytics across assets and facilities. It ensures industry experts spend seconds, not hours, to find and understand the data they need. Digital twins can codify domain expertise, continuously enrich your industrial data foundation, and leverage new technologies for better data capture and safer operations. And an open, extensible digital twin allows your organization to work with preferred industry standard tooling and visualization software, along with trusted partners.

By bringing all data into a single repository and contextualizing it, companies can rapidly build and deploy APM solutions to improve the performance of their assets and the people who operate them.

Organizations investing in an open industrial digital twin prove value from day one as they take advantage of out-of-the-box solutions and capabilities, connectors to existing systems, and an ecosystem of partners. Digital twins provide the means and the confidence to achieve the goal of an autonomous, self-improving industrial future.



Generative AI Copilots for Asset Performance Management

Rapidly maturing generative AI technologies are poised to deliver a step change in how humans process complex information or situations. While subject matter experts (SMEs) and operators are getting used to new data infrastructure that helps them 'Google' complex industrial data, LLMs such as Open AI's ChatGPT will further simplify the

human-to-data interface: the dawn of industrial copilots is upon us.

So what are the implications of this technology on legacy asset performance management processes and in live operational context?

Generative AI copilot experiences, powered by a robust data foundation, will transform how we work in industry.

Example use cases for generative AI copilots in Asset Performance Management

As an operator in the field, I don't have the time to double-check multiple sources (data and people) to gain context.

I need quick access to trustworthy data so I can understand real world conditions and take the appropriate actions to improve production ASAP.



As a maintenance manager, I am always trying to proactively plan, but I have little insight into operations.

Priorities can shift rapidly, and it is a painful process to shift resources to execute new plans while minimizing impact to scheduled activities.



It's very hard to understand the condition across many different equipment as data and analysis are captured in disjointed systems and some assets have very little data.

Oftentimes I need more complete information to take timely actions...



Workers in the field, either in the context of performing a maintenance activity (e.g., replacing a pump) or doing an operator round (e.g., an area check), often need to access documentation about the equipment they are working with. Based on what they are doing, this could be a start-up procedure for the pump you're working with or something as narrow as the pump's design temperature from the OEM.

Traditionally, this information would have to be printed before leaving the office and physically taken into the field. However, as often happens with fieldwork, discovering the need for new information requires another trip to the office and additional pieces of paper.

Now let's look at the future of APM.

By photographing the equipment they inspect, a generative AI-powered copilot system utilizes Optical Character Recognition (OCR) and image segmentation technologies to identify the specific equipment. This breakthrough feature saves valuable time and eliminates manual data entry, ensuring a seamless and intuitive user experience.

A copilot automatically analyzes the photos, highlighting any differences or anomalies that may have emerged over time. This automation enables field workers to quickly identify changes, such as the emergence of corrosion or other visual anomalies, facilitating proactive maintenance and ensuring the integrity of critical assets.

By comparing current equipment images with historical references, the copilot provides field

workers with a comprehensive and holistic view of the equipment's condition. This capability empowers field workers to make accurate assessments, take appropriate action, and address potential issues before they escalate. The seamless integration of image comparison technology streamlines the inspection process, reducing the time and effort required to identify equipment anomalies manually.

Copilots also make bringing up relevant information for any equipment incredibly simple. By tapping into the vast repository of industrial data stored in the open industrial digital twin, field workers, maintenance engineers, and reliability engineers have access to comprehensive and contextualized information right at their fingertips—whether in the office, control room, or field.



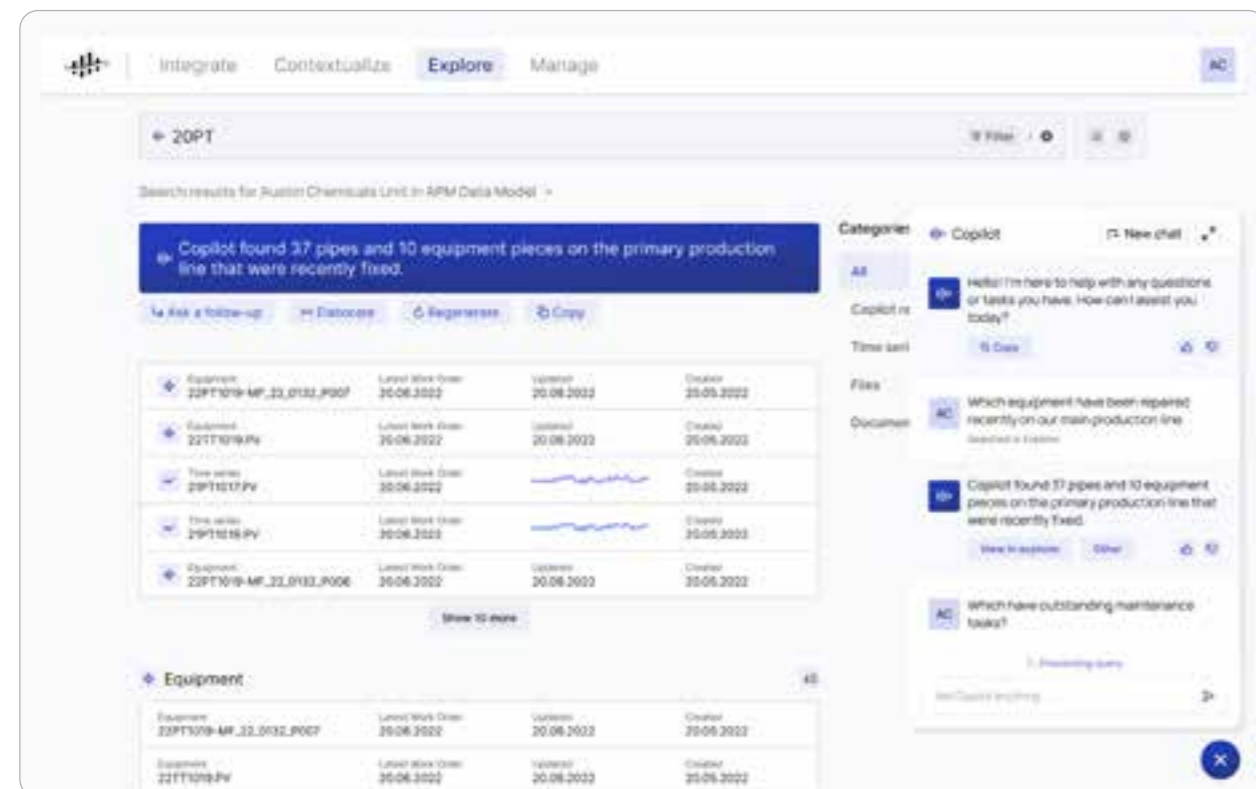
Such simple access enables SMEs to make informed decisions and perform accurate inspections without requiring extensive manual searching or reference materials.

Additionally, field workers can interact with the equipment using natural language through speech-to-text or text-based queries. With the power of generative AI, workers can ask questions such as the last pressure test date, the upcoming paint job schedule, or the maximum operating pressure of the equipment. The system intelligently analyzes the queries and retrieves the relevant information, providing immediate answers to field workers' inquiries.

In addition to information retrieval, these copilot experiences enable users to fill in work reports using natural language effortlessly. The copilot will automatically populate the relevant fields in the report templates, whether inputting a report manually or utilizing speech-to-text capabilities. This seamless integration of AI technology eliminates the need for manual data entry and ensures accurate and efficient reporting, enhancing the overall productivity of field workers.

By combining the power of image recognition, natural language processing, and seamless integration with the industrial knowledge graph, SMEs can perform their APM-related tasks more efficiently, with improved accuracy and reduced administrative burden.

Automated operator rounds performed by robots, drones, and other new technologies will play a crucial role—freeing up valuable time for SMEs from performing simple routine tasks and reducing their time in dangerous areas.



It's critical to find the right balance between investments in hardware automation (e.g., robots), software automation (e.g., AI and machine learning), and the skills of your human workforce. It's not about replacing all people with robots; rather, it's about finding the appropriate balance to draw upon the unique capabilities that each brings to the conversation.

With ChatGPT serving as the translator, people can do powerful work with robots and drones that would previously have been in the realm of developers just by typing in their intent. From the safety and comfort of their desks, an operator can troubleshoot an issue with their copilot, which

can simultaneously deploy a robot into the field for equipment inspection or repair—no need to interface with multiple applications or write a single line of code. Automation and remote work becomes as simple as typing a sentence.

Once again, it's not about replacing the human workflow 100%, but rather making the 80% of the work that solves 80% of the most frequent and valuable challenges much simpler.

“If humanity continues the chosen course of movement, it is apparent that the value of humans themselves will constantly and steadily decline. In many areas humans may someday become completely unnecessary to each other, and most importantly, machines will not need them either.”⁶

As we start to empower SMEs in asset-heavy industries with this kind of intelligent tooling, you begin to believe that we could be—at long last—at the tipping point of superhuman augmentation. What happens when people with decades of domain expertise and real-world equipment knowledge can self-serve and perform orders of magnitude more efficiently? That could very well be the inflection point of a new industrial revolution.

⁶ Gartner, Maverick Research: Preservation of Human Functions in the Age of AI. 8 May 2023. ID G00776866 →



The Next Phase of Digital Industrial Platforms

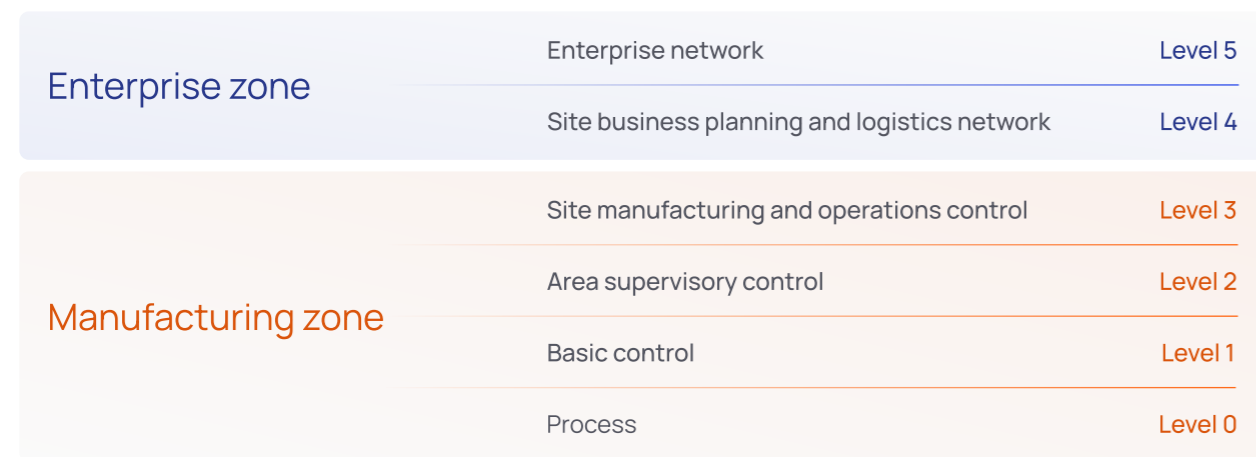


Whether you are a digital maverick or about to become one, now is the best time to work in digital transformation. Seriously.

We've come a long way since the early years of industrial IoT data management. Pioneered by operations technology incumbents focusing on on-premise hardware and software solutions for Levels 0-3 (see the illustration below), later established as first 'standard' for the platform approach defined by Gartner's Industrial IoT Platform Magic Quadrant in 2018, and most recently brought to the cloud era by numerous IIoT Platform look-alikes in a myriad

of variations of digital twin platforms. However, we remain very early on our journey. As a result, we commonly pay the cost premium of complexity and immaturity at every turn.

The Purdue reference model



The value of digitalization is well understood, even well documented. Scaling the value across the industrial enterprise, however, remains challenging. Data remains siloed across hundreds of on-premise systems, and public

cloud + consultants-led digitalization platform initiatives have proven much more challenging than anticipated. Great—or even good—user experience remains rare.

And Just Like That, the iPhone Moment Arrived

Generative AI is set to transform every class of software application and platform. There is now a clear market gravity towards a new industrial data ecosystem model that supports the whole range of cloud data and AI workloads, from operational and transactional data to exploratory data science to production data warehousing (exemplified by the rise of cloud data historians). Market gravity that places the transformative power of generative AI at the core of this emergent new industry cloud paradigm—one that makes intuitive user experience a first-class citizen.

“Using the traditional approach is too costly. The solution must work as part of this emerging enterprise platform and at an intuitive level for those using it.”

The quantified economic value



As we pay witness to the iPhone era of digitalization, we foresee generative AI introducing three intertwined transformations for industry:

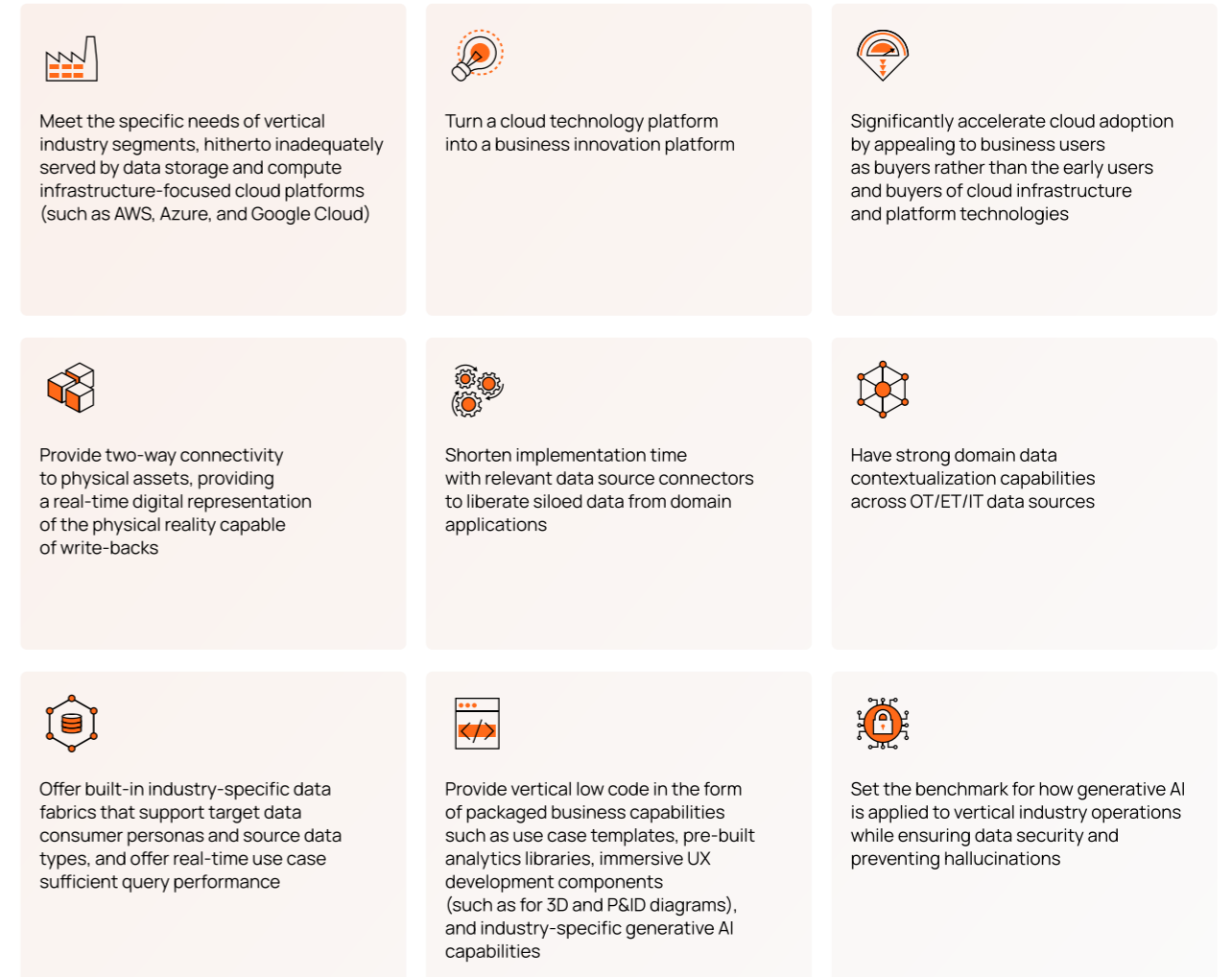
1. **Data and analytics roles will no longer be a priority:** AI will become the primary influence in most business decisions, automating the intermediate human data and analytics steps to the same extent e-commerce has automated consumer retail and banking.
2. **Going from low code to no code for 90% of all use cases:** With large language models (LLMs) instantaneously translating between natural language inputs and all programming languages and towards all APIs, everyone can finally become a software developer—or rather enjoy the benefit of their own industrial CoPilot as their personal software developer.
3. **Vertical software (so-called industry clouds) will replace more than half of all functional enterprise software,** providing industry-focused, cross-data source value traditionally achieved by extensive customization of general-purpose software with costly SI projects.

In fact, the catalyst role of generative AI accelerates the emergence of industrial clouds by increasing competition and innovation, as much as tangibly redefining the art of the possible—even in fields considered impossible to permeate by such new technology.

We will see more change over the next 18 months in the diverse digital transformation platform market (consisting of industrial IoT platforms, digital twin platforms, and more) than we have seen during the lifetime of Gartner’s Industrial IoT Platform Magic Quadrant from 2018 to 2023.



Industrial Cloud Platforms



Introducing the Industrial Cloud Powered by Generative AI

For most enterprises, by now, the cloud has disrupted conventional data centers by providing more elastic storage and compute services and improved mobility. What lies ahead is *the cloud as a true business disruptor*. Gartner expects that, by 2027, more than 50% of enterprises will use industry cloud platforms to accelerate their business initiatives.

So what is an industrial cloud platform?

An industrial cloud platform—or industry cloud platform more generally—offers industry-relevant, direct business outcomes by packaging IaaS, PaaS, and SaaS services around a robust data integration and contextualization backbone (commonly referred to as an industry data fabric) together with libraries of composable business capabilities such as use case templates, domain and solution data models, pre-built analytics schemas and dashboards, and industry-specific generative AI capabilities.

An industrial cloud platform is an end-to-end, integrated set of tools to simplify and reduce the time to production of applications for all data use cases. It reduces the cost of implementation, integration, maintenance, and management.

An industrial cloud platform is built around an industry-specific data fabric—simple access to complex industrial data from one API—with

common governance, metadata management, and unified access management, with a set of services accessible by the business user.

An industrial cloud platform applies safe, secure, and hallucination-free generative AI to offer an ultra-intuitive data and user experience for many vertical industry end users on the same unified collaborative platform.



An Aside: 6 Myths of Building Your Own Industrial Cloud

Now is a good time to talk about the alternative to using an industrial cloud SaaS; DIY (do-it-yourself). This is what most organizations have been doing on-premises for years. In fact, early adopters of cloud also had to resort to DIY in many places as proven SaaS alternatives did not yet exist.

The problems with DIY are many. Many of us know the amount of time and cost that goes into integration of disparate tools, or assembling from a variety of PaaS.

In addition, the level of technical expertise required to perform the integration is higher than implementing an industrial cloud SaaS.



What about the business value delivered by DIY? Would you be reading this if that value was a strong triple-digit percentage? As for Cognite Data Fusion® customer value, an independent study by Forrester recently concluded 400% ROI delivered.

Finally, DIY requires IT to be involved at every step, including after the project is finished, due to ongoing maintenance and support of the integrated tools. Therefore, we believe that although DIY is an alternative, it is far more complex and costly. And while generative AI is making using software easier, it is in fact increasing the complexity of developing enterprise solutions to be secure and hallucination free.

Let us examine common arguments used for going with DIY vs SaaS.

1. DIY is cheaper

Talking about cost alone is, of course, nonsense. Investing \$100K and not getting any return is a poor investment, whilst investing \$2M for a \$8M ROI is a solid financial investment.

Cost is also not the license cost paid to cloud and/or SaaS partners, but the Total Cost of Ownership (TCO) including the 'mortar' (people) needed to put together and maintain the data and digital platform. Unlike with normal business SaaS, the ongoing cost of the 'mortar' often far exceeds the software license cost for DIY platforms. This 'mortar' is, of course, part of the true cost.

It is not hard to find very large (100+ FTEs of 'mortar' alone) DIY platforms across industries, making the cost side rather high. Are the economies 'cheaper' with commercial SaaS that does not require similarly persistently high volumes of 'mortar'? This is very, very unlikely. So unlikely in fact, that finding reasons to not choose SaaS has always focused more on points two to seven on the list. In the more austere times ahead, this alone should speak volumes.

2. DIY is the only way to avoid vendor lock-in

One of the most ironic arguments around is that by going all-in with one cloud and DIYing everything using that cloud's proprietary services (as multi-cloud DIY is exponentially harder), this platform strategy will protect against vendor lock-in. The truth is, your enterprise is already multi-cloud, and will be more so going forward. Cisco Systems Inc.'s 2022 Global Hybrid Cloud Trends survey found that 82% of IT leaders have adopted a hybrid cloud while just 8% use only a single public cloud platform.

Same applies to your SaaS choices. Working with select specialists gives you access to relative strengths of different partners. It allows you to compensate for lack of a specific capability by one, with that of another. It allows you to maintain your flexibility of choice. If categorically avoids waking up locked in to one costly, lower performance solution with a very high TCO.

Openness, focus on industry standards for data exchange and data models, and placing partner ecosystems—including frenemy ecosystems—at the center of customer value delivery are not



unique to clouds. These are normative business hygiene for all of 2020s enterprise SaaS. Choosing SaaS to get to business value fast does not result in lock-in. Finding yourself with a high TCO solution that does not meet business needs on the other hand is lock-in of the worst kind.

If, of course, you fear getting locked into a SaaS vendor because your business users love using their solution over DIY alternatives, then you should jump to points four to six.

3. DIY is necessary to differentiate

Anchored in competitive strategy, the thesis that to outperform peers one has to be differentiated is true. Yet for manufacturing or energy companies to pursue meaningful differentiation by building proprietary data and digital platforms is missing the forest for the trees.

Building a proprietary data and digital platform is only meaningfully differentiating if it has a positive impact on the customer experience—in this case, how easily can SMEs use the platform to create business value. For anything technology

related, this more and more means positive product user experience (UX) differentiation, not technical features. Getting DIY UX to stand out above SaaS alternatives who are able to learn and iterate based on feedback from hundreds of customers—whilst your teams learn only from your users—is a moonshot.

The world is full of highly differentiating industrials using the same ERP (SAP) and the same CRM (Salesforce) for example. In fact, by leaving the tools of the trade to specialists (SaaS partners who only exist to provide these solutions), you can differentiate on what your customers actually value: on your customer service, your distribution, your ESG, your business model, and beyond. As for cost efficiency as a differentiator, go back to point one above.

4. DIY is the only enterprise-grade path

Enterprise SaaS is every bit as enterprise-grade as the first-party public cloud services used to develop a custom-made data and digital platform. The only difference is that whilst enterprise SaaS comes with comprehensive SLA and premium support options, with DIY, you are responsible for your own SLA and support. This may not sound like a seismic difference in the early phase of a DIY journey, yet when it comes to supporting thousands of users, having a strong user community and documentation—all table stakes with enterprise SaaS—warrant serious consideration. Enterprise grade means enterprise scale usage ready.

Moreover, enterprise SaaS comes with coherent, well-researched, painstakingly tested and iterated, business-user-persona-based UX. Whilst perhaps not yet the first on the enterprise-grade checklist, cutting-edge 'maverick' CIOs already place UX first for good reason.

And, of course, all enterprise SaaS comes with enterprise grade security. With all likelihood, your enterprise's customer data as well as employee data is already on an enterprise SaaS product—are your equipment time series and events really more confidential?

5. DIY gives me job security

The best means of creating job security and establishing value for your department and your team is to create business impact aligned with your company's strategic priorities and values. Delivering real business value to your business operations is the most direct, measurable, and visible way of doing so.

It is also worth noting that your business operations do not care what the brand of the tooling is. They only care that it makes their day-to-day lives easier, allows them to have better situational awareness on events in real time, and to make better business decisions. They really do not care if it is DIY built or SaaS or hybrid. They do care that it is fit for purpose, easy to use, and that it makes their life easier, not more difficult.

In different times, starting a promising, multi-year, three-figure (\$M) strategic DIY platform project was aspirational. Today, in different times and with the majority of those three-figures spent without much ROI to show for, focusing on here-and-now value to business is the best job security plan there is.

6. DIY allows me to play a strategic role

Meeting business needs today—and anticipating business needs of tomorrow—is strategic. Finding agile solutions to business needs that deliver value quickly; are easy for business users to adopt and scale; are open, interoperable, and extendable; and are flexible to onboard, maintain, and offboard if needed, is strategic.

Focusing on customer experience is strategic. Being open to new ideas, and being the one to introduce innovation (see industry cloud platforms in point five above) is strategic.

With all likelihood, your company is not turning into a software platform company, but is and will remain an energy or manufacturing company. Continuing to build and maintain the behind-the-scenes, bottom-of-the-stack platform services on which to build business value and differentiation is a necessary evil. Strategic it is not.

So turn a new clean page and think strategically about what would be best for your business regardless of sunk costs. That too is strategic.

From Optimized Data Storage and DataOps to Data Discovery and Composable Business Solutions

After years of costly DIY and 'significant assembly required' data management solutions—that even when painstakingly assembled, feel antiquated before even operationalized at any scale—there is strong argument for unified, easy-to-use, pre-

integrated set of tools to effectively manage and control data assets with minimal integration and setup costs, and which place the business user as the primary data beneficiary rather than the IT department.

Having led the way on Industrial DataOps platforms, we can summarize the evolution to industrial cloud powered by generative AI as:

From Industrial DataOps to Industrial Cloud powered by Generative AI



Generative AI is redefining low-code application development. It enables business solutions that were only dreams until now. Using a workbench-style UI—such as Cognite Data Fusion's Industrial Canvas—business users can create applications, both analytic and workflow (including with source system write-back using the data fabric), based on no-code user experience.

Industrial Canvas makes cross-data source insights available to everyone, at all levels of the organization, to easily build specific use case applications. This decreases time spent searching for data and increases time spent collaborating, accelerating high-quality business decisions by 90%.



Industrial Canvas overcomes the challenges of other single pane of glass solutions, which often over-promises capabilities and are too rigid with prescribed workflows, preventing users from working with the data how they choose.

Consider APM investments. The challenge is that APM solutions are often delivered in isolated workflows, limited in their ability to be tailored for unique site needs, and not structured to apply generative AI advancements. Additionally, manual data management and the need to customize applications across sites has prevented APM solutions from scaling. Many solutions still exist in on-prem silos, managed by different parts of the organization, and deployed to solve one or two narrow business problems.

Organizations need more composability in their solutions to extend asset life and maximize labor productivity, reduce cost and time of maintenance, increase equipment performance, and minimize unplanned events. Such value can only be delivered by empowering their people to make higher-quality decisions by reducing the efforts required for SMEs and field engineers to create self-service insights from data.



Industrial Organizations Require SaaS Environments to Scale Operational Impact

While the number of companies leveraging the speed and scalability of SaaS solutions is growing, many industrial organizations still have an aversion to any solution that will not run on-premises or in their private cloud.

In many cases these organizations believe that on-prem or private cloud environments are the only acceptable ways to securely store sensitive data, meet regulatory requirements, maintain control of their data, and avoid vendor lock-in. These concerns are valid, and a SaaS solution must fully address each concern while delivering faster time-to-value and solution scalability.

Before delving into how industrial companies can benefit from embracing SaaS, let's begin by defining the commonly accepted benefits and shortcomings of on-prem, private clouds, SaaS, and public clouds.

On premises (on-prem)

In an on-prem environment, organizations own and maintain the physical infrastructure (including storage, data servers, networking equipment, etc.) hardware, and software required to host and operate their systems. Organizations are responsible for costs to manage all aspects of the IT environment, including security, backups, upgrades, installations and more. On-prem is valued most for critical use cases such as process control where network interruptions do not impact performance. On-prem infrastructure limits scalability, flexibility, and innovation. Managing and maintaining on-prem solutions require additional investments in infrastructure and resources.

Private cloud

Private cloud is a cloud environment dedicated to a single organization, providing them with exclusive use and control over the cloud infrastructure and resources. It offers enhanced control, security, and customization options but usually requires high upfront costs and ongoing maintenance. Private cloud is valued for the control, reduced hardware management, and increased scalability over on-prem solutions. However, the infrastructure still requires dedicated IT teams to manage and allocate resources, including software and security upgrades and updates.

Public cloud

Public cloud environment infrastructure and services are shared across multiple users/ organizations. The services and resources are accessible to the public over the internet, and customers can access and utilize them on-demand. Concerns around public clouds exist around security, especially related to storing of sensitive data outside of an industrial organization's infrastructure, and time-to-value, where initial projects can take six-plus months to deliver business value. Public clouds are valued for their scalability, cost-efficiency, and a wide range of pre-built services. The IT environment, software upgrades and security updates of underlying services are handled by the cloud provider.

Software-as-a-Service (SaaS)

SaaS is a software delivery model in which software is accessed through the internet/public cloud, eliminating the need for local installations. Under the SaaS model, the software is stored on remote servers, continuously managed and updated by the service provider, and easily accessed by customers through web browsers, mobile applications, and application programming interfaces (APIs).

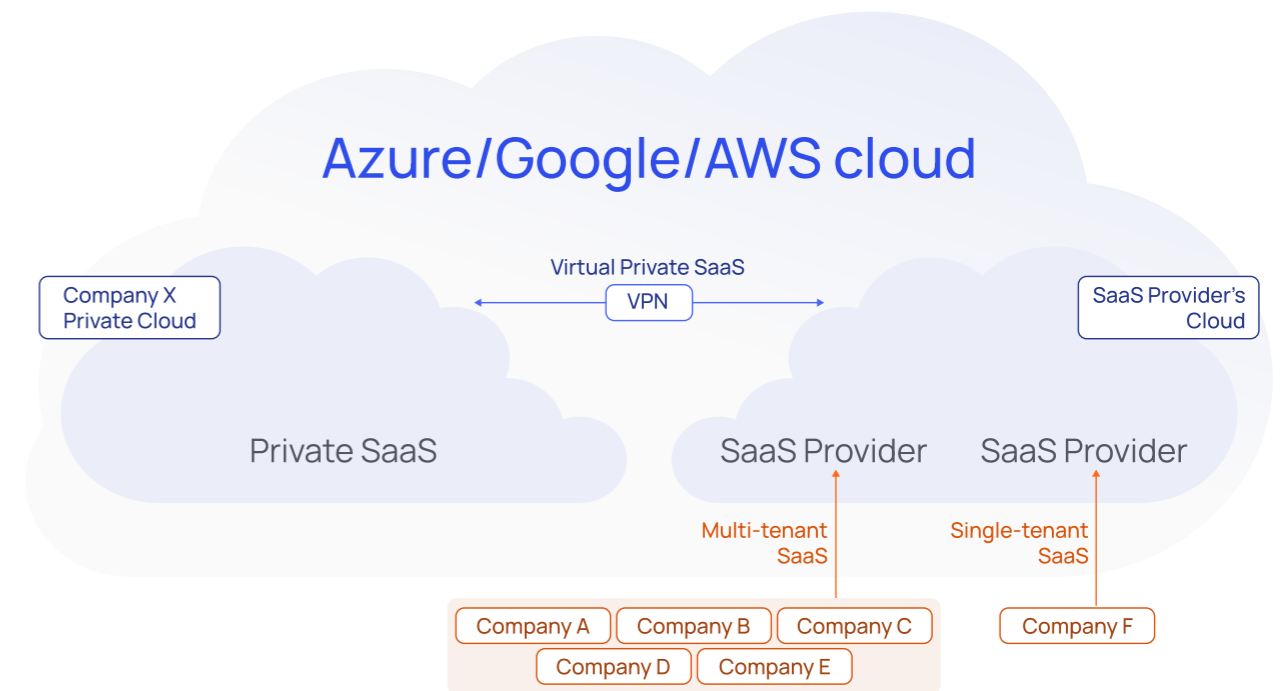
It's important to note that SaaS can be deployed in multiple ways:

Multi-tenant SaaS	Customers share the same resources with their data kept totally separate from each other.
Single-tenant SaaS	Customers do not share resources and have their own dedicated cluster with their own software instance.
Virtual private SaaS	Customers do not share resources and have their own dedicated cluster with a secure connection (private IP address) that is not accessible/ findable via public internet.
Private SaaS	Software solutions are installed inside of a customer's private tenant, giving customers increased responsibility and control over deployment, security, maintenance, and data.

When evaluating different SaaS environments, many have the initial reaction to deploy these solutions within their own private tenant. Deploying inside of one's own tenant passes responsibility and control over deployment, security, maintenance, and data to the industrial organization. While this may be tempting, it is essential to consider the following:

- **Deployment costs:** While counterintuitive, the more control customers want to own the environment, the higher the cost to the SaaS provider. Depending on the complexity of the solution, a private SaaS solution may be 10-20X more/year than using a multi-tenant.

- **Vendor guarantees their security standards:** SaaS providers practice comprehensive security and privacy measures and conduct thousands of security updates daily. Managing such a level of security within a private SaaS will require a dedicated security team. Moreover, vendors are required to undergo testing by third parties to obtain major security compliance certifications (e.g. ISO 27001, SOC 2 Type 2, etc.). The majority of industrial companies have not received or maintained these certifications.



- **Inaccessible from the public internet:** This can be a requirement for some critical industries or sensitive data. Both virtual private and private SaaS make data inaccessible from the public internet. The virtual private SaaS uses a virtual private network (VPN), granting it a local IP address, and the SaaS solution runs in the same data center as the customer's tenant, making it indistinguishable from their own network.
- **Software updates maintained by service providers:** All software maintenance (updates, upgrades, bug fixes, etc.) is handled by SaaS providers. Yet, if an organization chooses private SaaS deployment, it either needs to manage isolated software installations, upgrades, and necessary infrastructure expansions on its own or rely on the SaaS provider or partner, likely at an additional cost.
- **Vendor guarantees SLA and user experience:** For providers to guarantee a service level agreement (SLA), they must have access to control the underlying services and infrastructure. Private SaaS deployments can be a blocker for providers to ensure their SLA is met. Additionally, because providers cannot control the resource allocation, they cannot guarantee an interactive user experience as other workloads running in the private cloud may impact performance.
- **Customer retains data ownership:** This must be true regardless of the chosen SaaS environment. An organization's data is always their own to share, copy, or delete as they see fit.
- **Data is never stored outside of the customer tenant:** The private SaaS is the only environment where this is fully guaranteed. Achieving this has significant tradeoffs regarding higher costs, effort to maintain security, SLA, and a delightful user experience.

SaaS environments supported by Cognite				
	Multi-tenant SaaS	Single tenant SaaS	Virtual private SaaS	Private SaaS
Deployment cost	Ⓢ	Ⓢ Ⓢ	Ⓢ Ⓢ Ⓢ	Ⓢ Ⓢ Ⓢ Ⓢ
Provider guarantees their security standards	Yes	Yes	Yes	No
Inaccessible from public internet	No	No	Yes	Yes
Software updates maintained by service provider	Yes	Yes	Yes	Yes*
Provider guarantees SLA and user experience	Yes	Yes	Yes	No
Customer retains data ownership	Yes	Yes	Yes	Yes
Data is never stored outside of the customer tenant	No	No	No	Yes

*Requires specialized access to the Private SaaS environment by the vendor

At SaaS companies, such as Cognite, there are hundreds of specialized services and people that ensure the environment is stable and highly performant for customers through a SaaS experience. Thus, for a typical industrial company to successfully manage a complex SaaS solution in their private tenant, consistent with a SaaS user experience, it would require a dedicated team solely focused on maintaining the software.

To provide the highest level of availability, security, and performance look for a SaaS environment that meets the following measures:

- Maintains staging environments that are 100% in sync with the production environment for proper testing and validation of updates. Environments are updated thousands of times per week to provide necessary improvements and optimization.

- Manages a single version of generally available software. All customers benefit from new releases, without having to migrate versions. Managing a single version reduces risk and increases the rate of innovation and new features, such as generative AI-powered search.
- Tenants are clean from other processes, data, and jobs with enforced governance to ensure that at any time users access information through Cognite Data Fusion®, they use the proper API channels, enforcing user access rights, security, scalability, etc.
- 24/7 monitoring and support using specialized deployment technology and operational management tools.

- Security tested and audited by third parties with compliance to ISO 27001, ISO 9001, SOC 2 Type 2, CCC+, World's first DNV certified digital twin. These certifications are guaranteed when operating in a Cognite multi-tenant, single tenant, or virtual private tenant. Cognite cannot guarantee security standards if a deployment were to run in a customer's private tenant/cloud.
- Support for specific regulatory and industry requirements including: NIST CSF, IEC 62443.2-4, IEC 62443.3-2, IEC 62443.3-3, IEC 62443.4-1, CMMC, FIPs, NERC CIP v.5, GxP.

Additionally, you want a SaaS solution that can be offered as a multi-tenant, single tenant, or virtual private tenant, all depending on the needs and requirements of the end users. For the most critical industries and most sensitive data, for example, our customers use Cognite Data Fusion® via a virtual private tenant. Cognite's virtual private SaaS environment combines the security and privacy of a private cloud with scalability and flexibility of single and multi-tenant environments.

In this environment, Cognite Data Fusion® runs in the same data center as a customer's private tenant, eliminating any regional data residency requirement concerns. Configured within the virtual private local network, it is granted a local IP address, exclusively to your company and unreachable from the public internet. All data ownership is retained by the customer and only accessible through documented APIs and connectors. The customer always retains ownership of the data.



Our customers trust us not only because we are compliant with the multi-tenant architecture and follow the best practices from hyperscalers like Microsoft, Salesforce, Google, Adobe, and others, but also because Cognite has domain expertise in energy, manufacturing and power and renewables. We understand the importance of providing a secure and reliable SaaS environment that is flexible, scalable and quickly adapts to growing business needs.

Tools for the Digital Maverick



Nobody wants to be in the situation where, after months of vendor meetings, internal alignment, tech reviews, and security checks for a new software purchase, your executive sponsor says “I’m just not buying it...” But that’s the reality for many complex enterprise purchases today, even with board-approved budgets opening up for AI software and an accelerating pace of procurement due to the complexity of decision-making involved.

Today’s digital transformation and operations executives in chemical production, refining, and energy are under increasing pressure to deliver more productivity, reliability, and safety with fewer human resources—but more software. According to recent Gartner Maverick Research, “by 2030, 75% of operational decisions will be made within an AI-enabled application or process,” demanding strategic—and rapid—realignment around technology.

But embarking on mid-game data and technology investments can be quite complex and sensitive, especially given existing entrenched software (SAP, MAXIMO, AVEVA, etc), siloed data management practices, and general resistance to change. Is the overarching strategy clear and communicated? Who needs to be part of the buying center? Who’s personally invested in competing strategies? What impact will these decisions have on future teams and workflows?

In this chapter, we provide you with the tools you need to scope and plan your digital journey, as well as define and measure success in terms the enterprise will understand. With the following tools, you now have what you need to plant your flag as a digital maverick and design and implement an innovative, future-proofed digital solution that will enable your organization to transform the way you use data. A key area that can either kill or delay a digital transformation is when key strategic questions are not acknowledged or addressed by the right stakeholders in the buying center. Here are a few common ones we see often:

What’s the Honest State of Your Data and Digital Maturity for AI?

We get it—your board is pushing hard to invest in generative AI projects. But getting buy-in, approvals, and operational value/success becomes easier when you incorporate your larger data and analytics journey. Coupled with the previous questions around use cases, by honestly evaluating your current technical strengths and weaknesses around data, you can plot a more accurate roadmap with the right technical components to either quickly build the maturity or finesse the maturity needed to deploy AI at scale.

Have Use Cases and Values Been Defined and Mapped?

Business value must lead the way when it comes to complex digitization projects. Rather than ‘eating the elephant’ all at once, projects must be broken down into manageable use cases and prioritized based on strategic near or long-term value. Doing this with enough detail and a product roadmap-like mindset ensures that AI-enabling software purchases are viewed by your buying center as an investment, not a cost, making it more straightforward to ensure buy-in and justify funding.

Is Everyone Aligned to the Same Software Strategy?

Often, there are competing viewpoints on DIY vs SaaS vs a hybrid approach; IT wants to build, while the business wants to buy. Compounding this misalignment is the fact that sometimes there is no clear decision-maker. This is where your digital maverick must lead and evaluate what’s in the best interest of the business. Is it more important to move quickly and innovate faster than your peers? Or is it more important to build from scratch and sacrifice speed for the sake of tech stack completeness.

To help you drive alignment across your organization, within this chapter you will find:

- **The Digital Maturity Diagnostic:** To assess the current state or track the progress of your digital maturity.
- **The Digital Roadmap Template:** To align priorities and get buy-in for a broader digitalization initiative.
- **The SaaS ROI Calculator:** To quantify the value of a SaaS vs DIY for your organization.
- **The Value Framework:** To increase your chances of obtaining and maintaining internal buy-in, secure a budget, and successfully articulate ROI across the business.
- **The RFP Guide:** To ensure you account for all critical capabilities/functionality required for success.



Assessing if You Are Ready for Generative AI: The Digital Maturity Diagnostic

Building organizations in support of digital maturity can look different at a company level or within an individual line of business. However, high-maturity organizations share some common attributes, segmented by internal versus external factors (see: Factors Influencing Digital Maturity).

Generated after evaluating a number of organizations in asset-heavy industries and combining the findings with third-party research, these five key factors of digital maturity form a sophisticated framework for discussion and evaluation. Unlike other frameworks, this one includes information architecture and technical data layers; both critical—and often undervalued—indicators of digital maturity.

We can further segment these key factors into additional parameters that can be measured and weighed together to transform the concept of digital maturity into a diagnostic framework. Use this tool to assess the current state or track the progress of your digital maturity.



Digital maturity diagnostic framework

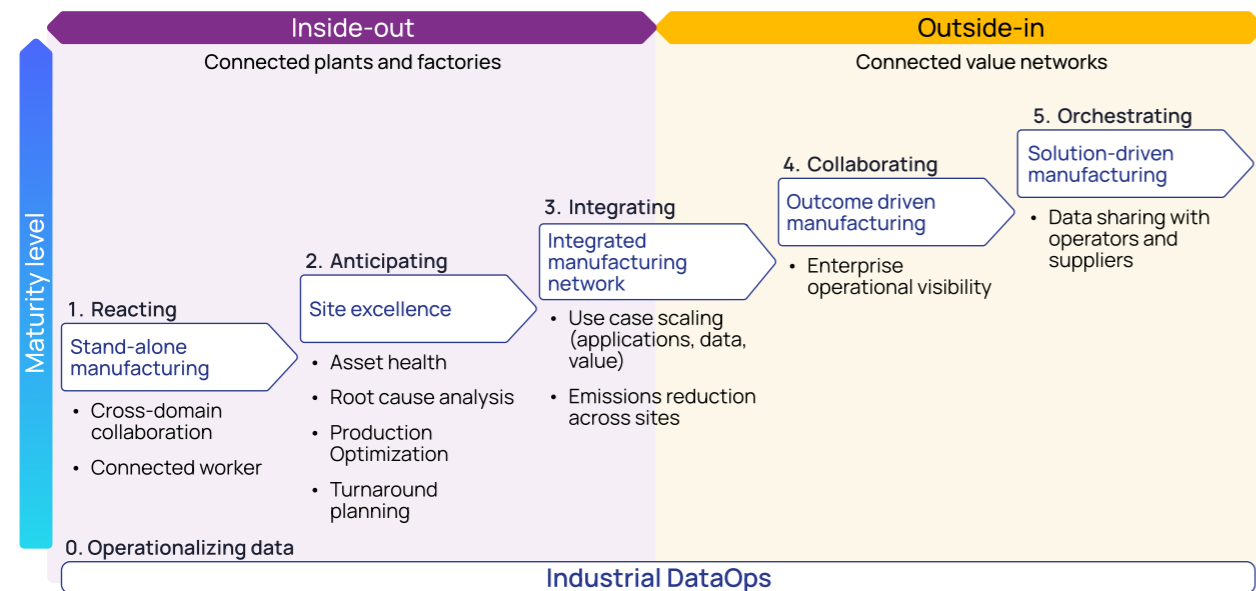
Category	Parameter	
Status of data	Enterprise architecture How integrated, flexible, and open is the enterprise OT, IT and engineering data?	Discovery What is the process for finding relevant data to a problem set?
	Operational technology & data How evolved, structured and managed is the process for maintaining OT source system data?	Standardization How well are data standards applied across sites and BUs?
Generative AI	AI relevance & use case execution Have you identified use cases where generative AI could add value to your business and teams?	Data leakage & trust What is your plan to manage data leakage/trust and access control when implementing generative AI on your private data?
	Industrial data context How will you minimize hallucinations? Have data quality issues been identified and addressed?	AI integration How easy would generative AI technology/models be to integrate with the existing systems?
People and mindset	Capabilities How advanced/capable is your digital organization (data scientists, developers, etc.)?	Adoption How will your subject matter experts (SMEs) in the field be empowered to use new data and generative AI tools?
	Implementation How do you plan to manage infrastructure implementation to integrate emerging AI technology into your operations?	Impact and value What is your plan for managing the impact of generative AI on workflows and employees?

By answering each question in the diagnostic on a scale of one to four (with one being “we have no plan” or “we are not confident” and four being “we have a detailed plan” or “we are very confident”), you can begin to establish where you are on the digital maturity spectrum and zero in on where to focus your efforts for maximum return.

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.

A very helpful reference point is Gartner's Maturity Model for Manufacturing Excellence. The five-stage maturity model is designed for manufacturing strategy leaders to assess their organization's current capabilities, articulate a plan for change and support the development of a

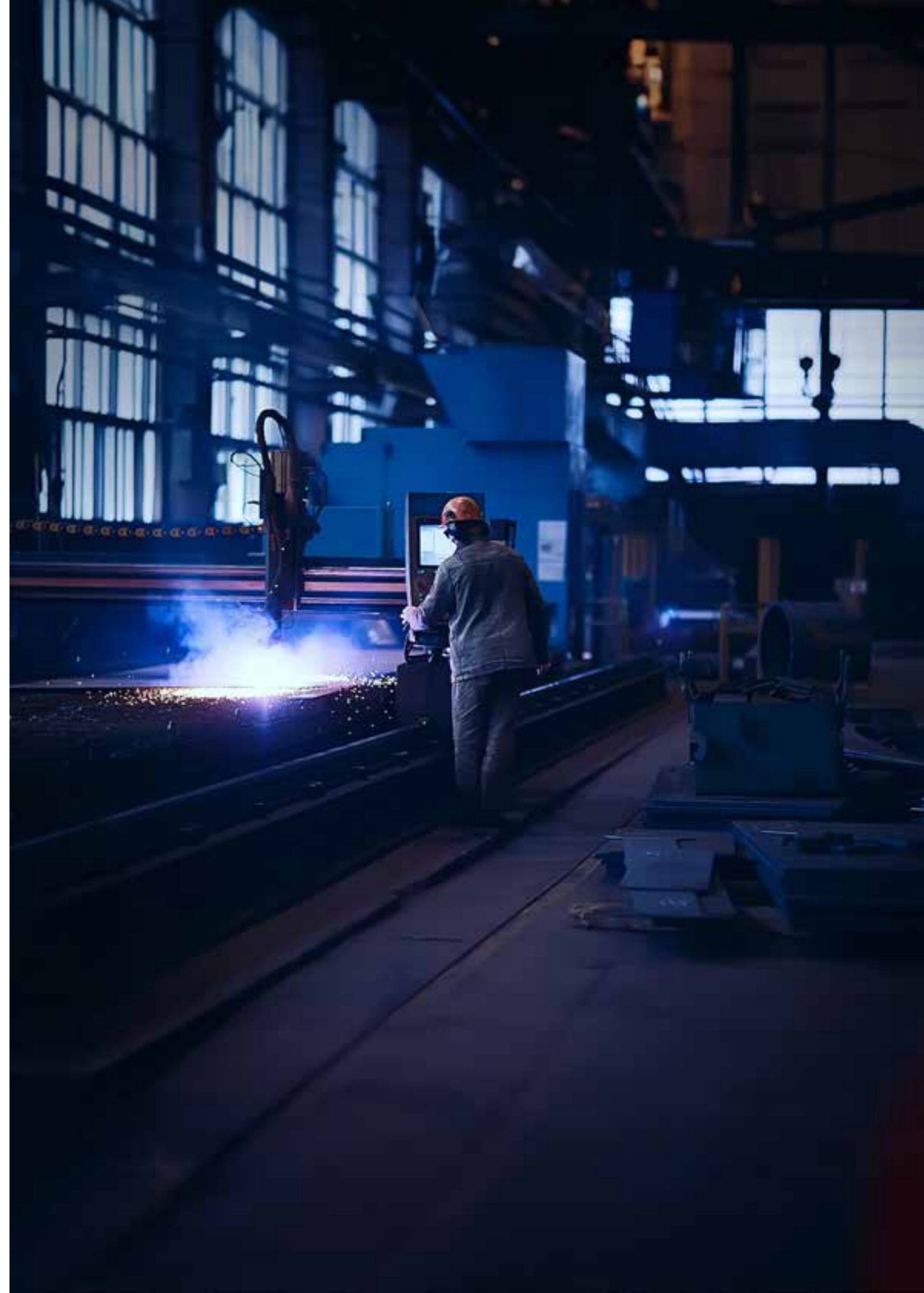
future-state vision for manufacturing operations. Cognite has taken this model and applied use case examples to each stage that the digital maverick can focus on to show how value can first be delivered at site level and then scaled out to the enterprise level for full value realization.



These tools are designed to be used together, first by taking the Digital Maturity Diagnostic and then by mapping your results to the corresponding stage within the Gartner Maturity Model. For

example, if you answered mostly ones in the Digital Maturity Diagnostic, focus first on building a strong data foundation to address use cases within the reacting stage of Gartner's maturity model.

Digital Maturity Diagnostic Result	Gartner Maturity Model Stage
1	Industrial DataOps / Reacting
2	Reacting / Anticipating
3	Integrating / Collaborating
4	Collaborating / Orchestrating

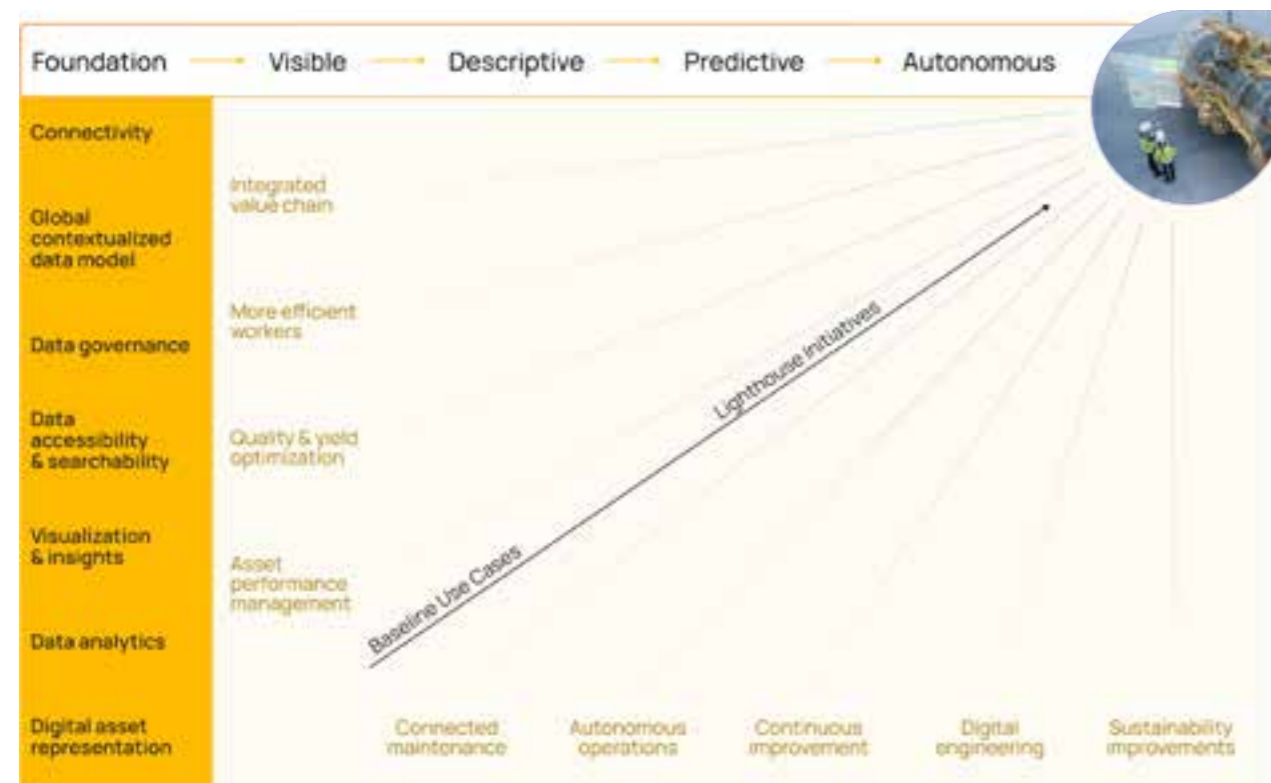


Prioritizing Industrial DataOps Use Cases: The Digital Roadmap Template

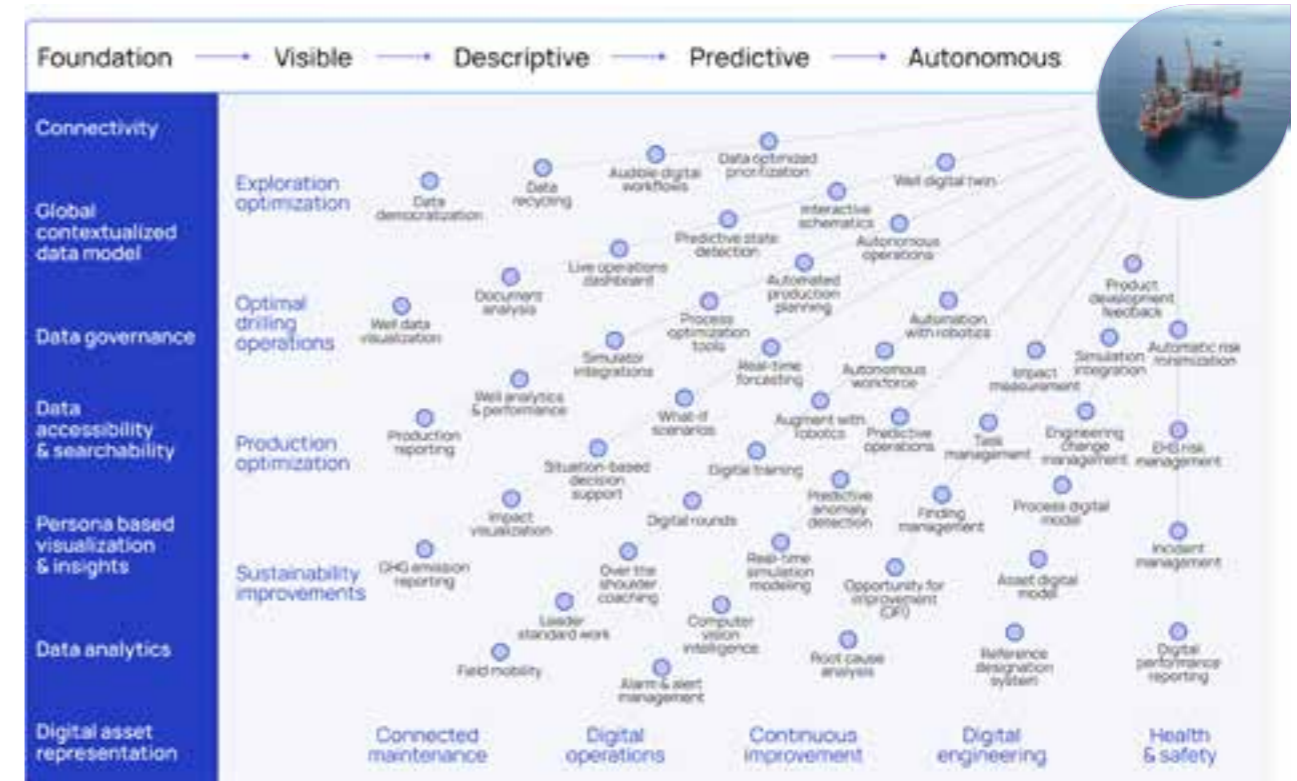
Historically, individual departments spearheaded digital initiatives, and each use case was considered a separate project. For example, at an oil and gas company, one team was looking into autonomous operations. At the same time, another team was focusing on real-time simulation modeling, and a third trying to optimize alert management. These are actually connected use cases, and they can build on one another to achieve more efficient success. At the same time, sustainability initiatives may be going on that would benefit from the data and dashboards the other teams are developing.

In fact, all digital initiatives are connected and benefit when built on top of the same accessible, trustworthy, and contextualized data foundation. Digital mavericks know this, but their challenge is connecting these disparate use cases into an actionable digital strategy.

This is where our Digital Roadmap Template comes in. Use this template to visualize your business strategy and make communicating with all stakeholders easier. This tool is critical for aligning priorities and getting buy-in for a broader digitalization initiative, not just a one-off use case.



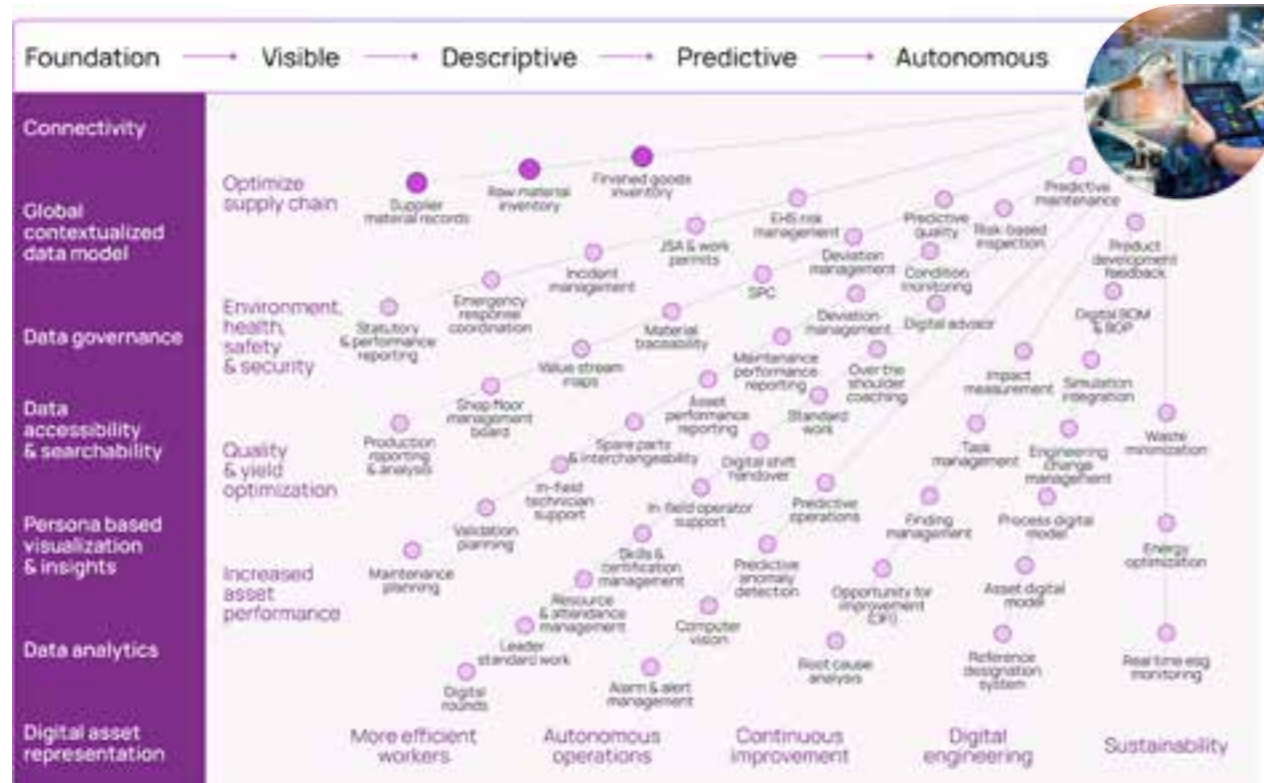
Example 1



Tips for Using the Digital Roadmap Template:

- The blue categories represent some of the most common focus areas across industry, but these should be tailored to your business. Examples of focus areas per industry are provided (see Fig. X).
- For each focus area, you will have several use cases mapped from baseline to lighthouse, visualizing the need for a solid and unified data foundation first, and second, the order in which use cases must be prioritized to reach specific long-term objectives. Examples of completed roadmap per industry are provided for reference. (see: Fig X, Y, and Z).
- The gray box on the left applies to all industries and represents the required data foundation for any successful digitalization initiative. This foundation includes the pillars of Industrial DataOps that will accelerate digital maturity and enable your organization to deliver better digital products and realize more operational value at scale.
- The orange bar at the top applies to all industries and represents the digital maturity required to implement a specific use case successfully. Baseline use cases make you think, “why haven’t we done this yet?” and can be accomplished early in a digital maturity journey. Baseline uses cases also act as building blocks to achieving lighthouse initiatives, such as fully autonomous operations.

Example 2



To help you determine which focus areas and use cases may be most impactful to your organization, consider the following:

- What is your operational view today? How easily can your team find everything they need to make the best decisions?
- How much of your operation is manual vs. automated? How well is that working for you?
- How does your team excel in making decisions? Where could they use more support?
- How predictable is your production and production quality? What are the most common causes of disruption to production?
- How does your team go from an alert to a resolution? How long does that take?
- What types of tools would your team build for themselves to help them make better day-to-day decisions?
- What processes would you improve if you could access high-quality data to automate them?

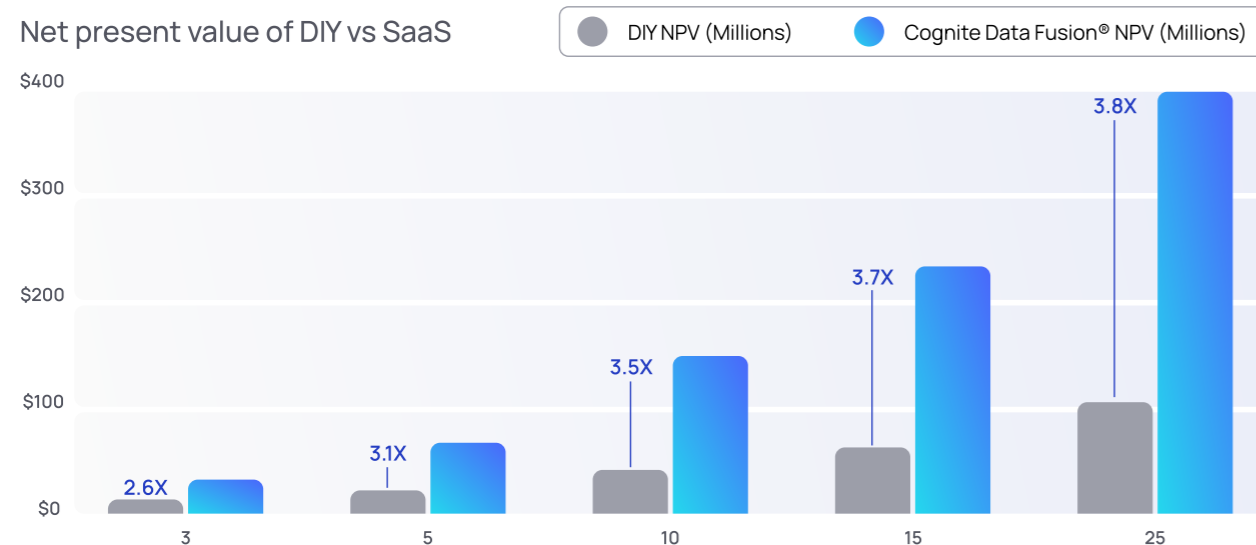
Quantifying the Value of SaaS vs DIY: The Value Calculator

Now that you have defined a long list of potential use cases and prioritized those use cases with a digital roadmap, you can build your case for a DIY vs SaaS vs a hybrid software strategy.

While pure-DIY approaches can be justified—especially in the early phases of emerging technology markets where SaaS vendors may be too small and lacking maturity—embarking on a DIY journey is suitable only for organizations with well-resourced IT teams and a high tolerance for risk. For most organizations, the benefits of deploying a DataOps foundation via Cognite Data Fusion®—a faster time-to-value, greater scalability, and lower maintaining costs—make more financial and competitive sense.

Across Cognite's customers we have seen that the Net Present Value (NPV) over a five-year period for Cognite Data Fusion® can be almost four times higher compared to DIY, underscoring the significant financial benefits of adopting a SaaS approach. A full-scale, enterprise-wide Industrial DataOps foundation has the potential to generate hundreds of millions of dollars in NPV, with Cognite customers achieving up to \$300-500m in value.

The below value graph will help you understand the total value generated based on the number of use cases deployed and give an indication of the value you can achieve to help you drive internal excitement and garner support to conduct a future value assessment based on your unique use cases and situation.



Defining the Impact of Industrial DataOps: The Value Framework

It is essential to have a clear framework for measuring returns on digital investments. Value frameworks are a necessary tool for executives to assess their business. While seemingly high-level, if executives cannot map their digital investments directly to one of the below value streams, they will ask themselves whether their investment is actually generating real value.

Value stream frameworks are a representation of how a company generates shareholder value and which levers it can use to do so. These are specific to industries and business and should be tailored for max impact:

- **Level 1** describes the top-level actions/metrics which directly affect a company's shareholder value; Support Revenue Growth, Reduce Operating Expenditures, Increase Asset Efficiency, and ESG.
- **Level 2** is the high-level business challenges that are tangible and measurable but are typically indirectly improved by organizational actions.
- **Level 3** is the value levels that are tangible, measurable, and actionable. Projects and the implementation of solutions directly impact these metrics and drive improvement in the higher-level categories.



This framework increases your chances of obtaining and maintaining internal buy-in, securing a budget, and successfully articulating ROI across the business. To align your digital initiatives within this framework, you'll want to consider how your use case(s) impact current challenges and help avoid adverse financial consequences. It is also essential to consider your ideal future state, the positive economic and operational impact, and how you will measure success.

The below questions are designed to help you evaluate your use case through a value-framework lens. While not comprehensive, these questions will guide you to a more value-centric mindset and enable you to lead productive discussions about digital strategy across the business:

- What would it look like for your team if there were no roadblocks to the information you need to make decisions? What would it mean for the business?
- In your desired future state, what positive outcomes do you see from taking advantage of de-siloed and contextualized data in your use case(s)?
- Who on your team and/or across the business would benefit most if cross-functional information was more accessible?
- Where would you focus the freed-up talent if it took less human power to deliver on your use case(s)?
- What would it mean to your bottom line if you could increase the speed of making a decision?
- What set of information would you include in your optimal KPI dashboard?
- How could better access to high-quality data improve production quantity and quality?
- How much of your budget covers scrap and waste across your sites/facilities? Which initiatives would you reallocate those funds towards if you reduced waste?
- How will better and faster visibility to scope one, two, and three emissions impact your business?



Finding Your Industrial DataOps Solution: The Request for Proposal Guide

A key challenge in implementing your own Industrial DataOps initiative is defining what capabilities your solution needs to support your business. The section below provides a guideline to build out your RFP and ensure you account for all critical capabilities/functionality required for success. This guideline will address key areas to consider and should be used as a starting point to tailor this to the needs of your organization.

Industrial DataOps RFP Guideline

Use Cases and Past Successes

- 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 client's 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 one to two use cases defined before any work begins and have a backlog of two to five use cases once success is achieved with initial use cases.



- How many existing customers do you have? Are there past successes you can share related to the clients industry?
- Does the proposed solution enable more effective asset management? Can you provide examples?
- What use cases have you delivered regarding unstructured data (video, 3D, etc.)?
- 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 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, scalability, and repeatability. Connectivity has two components—data extraction and application layer. Data extraction will ensure that you will be able to connect to both existing and future data sources while the application layer focuses on how the solution provider will support applications on top of the foundation to deliver use cases.

Foundation

- How does the proposed 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: Ideal solution should automate this process as much as possible or 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 this data to run an analytics model measuring performance.



- Does the proposed solution support templating? How can applied work be reused?

Expert tip: Templating 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, administrators, etc.?
- How would you describe the proposed solution's performance in regards to scalability?

Expert tip: As you expand beyond initial use cases, you will want a solution that is able to scale. Industrial DataOps should be able to address scale at both the site and enterprise level.

- How does the proposed solution process large data sets? How do you ensure the proposed solution can handle peak processing?
- How does the proposed solution support trending analysis of the data? How are trends visualized and reported?

- Can the proposed solution analyze trends in data quality and predict when metrics will exceed predefined thresholds?
- How does the proposed 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: While many solutions talk about openness, seeing examples of proven solutions with third-party vendors will provide confidence that you will be able to connect your disparate data sources.

- Is the front-end framework of the proposed solution built on open standards? How do you support open front-end frameworks?
- How does the proposed solution ensure that data is processed quickly and readily makes available time series data?
- Does the proposed solution require plugins such as Office, Flash, etc.?

Expert tip: Access to centralized, remote relevant-time data creates opportunities for many new use cases at both the site and enterprise level.

- Is the proposed solution able to ingest both tabular and graph-structured data without loss of information?
- When receiving asynchronous time series data, how does the proposed solution handle timestamping?
- Is the proposed solution able to handle data inserts, updates, and deletes by itself?
- Does the proposed solution support multiple modes of operation, such as batch and stream-based ingestion and in-memory versus persistent data storage?
- Does the proposed solution follow agile development principles and how do you ensure it is up to date on market trends and technical standards?
- How does the proposed solution support compression of data and metadata?
- Does the proposed solution report the source for a data point, event and time series, and 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?

Connectivity

- How does the proposed solution support integration with external systems and what are the requirements of such integrations?
- What integrations are pre-built and readily available for data extraction? For the application layer?

Expert tip: Pre-built 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.

- What are the client's possibilities to develop their own applications on top of the product?

Expert tip: Further assessment is needed when thinking about application development for data engineers and citizen data scientists. Proposed solution providers should have prebuilt 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 do you support? What connections are most common?
- What is the proposed solution's capability to access real-time data? What are the scalability limitations to this capability?



- Does the proposed solution have connectivity and native access to relational databases?
- Does the proposed solution 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 proposed solution support versioning for continuity so that the newest version, and the previous version of data pipelines remain supported? Can versions be rolled-back?
- Does the proposed solution support a layered and scalable REST API?
- Is the REST API stateless, enabling easy caching and no need for server-side state synchronization logic?
- Can underlying data be exported from the proposed solution as a CSV or XLSX? How does the proposed solution export data and metadata in standardized formats?
- Are there any limitations in the proposed solution's ability to extract historical data?





Generative AI

- How have you applied machine learning (ML) solutions to solve client use cases? Are there any use cases that utilize a hybrid AI (combination of physics and ML capabilities) solutions that you are able to share?
- How is generative AI incorporated into the proposed solution? Are AI capabilities built into the backend of the product? Is there a natural language copilot as part of the user interface?
- Can you provide details about the training data used to develop the AI model? What is your process for providing industrial context to generative AI solutions?
- How do you mitigate hallucinations within your generative AI solutions?
- How do you manage data leakage within your generative AI solutions?
- How do you manage trust and access control within your generative AI solutions?
- Is your generative AI solution compliant with relevant data protection and privacy regulations (e.g., GDPR, CCPA)?
- How frequently is the AI model updated and retrained to maintain its security and reliability?
- How does your solution handle intellectual property rights and content ownership?
- How have you applied generative AI solutions to solve client use cases? Are there any use cases that you are able to share?
- Are there any limitations or potential risks associated with using the generative AI solution?
- Can you share your generative AI roadmap and what are the most important near-term deliverables on this roadmap?

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 designed to meet the requirements of your existing environment.

- Can you describe the key components of the proposed solution and how 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 lake 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 proposed solution's ability to support real-time deployment?
- How does the proposed solution support horizontal and vertical scaling?
- How does the proposed solution offer high availability and how are failover procedures handled?
- How does the proposed solution support backup and recovery procedures?
- How does the proposed solution handle archiving?
- How do you support edge capabilities? Do you offer on-premises deployments?
- Is the proposed solution validated with the standards of W3C and HTML5 to enable browser independence?
- Does your proposed solution track the lineage of all data objects and code, showing upstream sources and downstream consumption?
- How does development occur in the proposed solution, introducing changes to its core components, adding extensions etc.?
- How is it possible to test reconfigurations, upgrades, and extensions to the proposed solution before it is 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 assess time to value and a high-level roadmap for implementation. The potential solution provider should provide the resources to ensure continued success. Successful implementations require both the right technology and the right support. This section is designed to provide insights to the expected support for your team and organization when adopting an Industrial DataOps solution.

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

- What maintenance and support do you offer during and after implementation?

Expert tip: 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 clients in-house experts to maximize the benefit from collaboration?

- How do you enable/support search in the proposed solution? Can you provide documentation?

Expert tip: This functionality will be very valuable to save time for data engineers and make data discoverable for citizen data scientists.

- How does the proposed 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 your client's organization?

Expert tip: Building competence within your organization is an important trait to digitally maturing. 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 IT departments often have developed their own security requirements for new software products and is necessary to consider when integrating IT and OT data.

- What is your company's strategy for penetration testing and third-party assessments?
- How does the proposed solution maintain an audit trail of all data manipulation?
- How does the proposed solution offer monitoring and statistics of backbone components?
- How do you ensure that the client has access to its own data in the proposed solution?
- How is high availability maintained for security, access, and governance of the proposed solution?
- How do you support revocation of access at both user and group level?
- When and how is data encrypted in the proposed solution?

- What is the proposed solution's capability with regard to access control? What is the granularity?
- Does the proposed solution 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 proposed solution support ISO, SOC 2 Type 2, NIST CSF, IEC, NERC CIP, or other relevant industry standards?
- How does the proposed solution track the chain of custody?



Usability

For products to be adopted by your organization, solutions must be easy to adopt by the users. Poor usability is a leading cause of poor product adoption. The potential solution provider also needs to support both data scientists and citizen data scientists to truly make data usable. In order to make data discoverable and usable for both of these users, the proposed solution must deliver simple access to data and provide intuitive, well-designed user interfaces that do not require strong coding backgrounds to leverage.

- 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.
- Are users able to create and edit their own dashboards to solve specific business needs, and are they able to share these dashboards to collaborate with their team or across teams?
- Do users see and feel the proposed solution responding in real time?
- How many concurrent users does the proposed solution support? And is the environment collaborative?

Expert tip: As your Industrial DataOps solution gains adoption, your organization should be striving to increase user adoption to enable use case development throughout many departments.



- How does the proposed solution handle error messaging? Are errors easily interpreted by users?
- How does the proposed solution allow users to 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, lack of service, and requests for new services or extensions to existing services?

Software Maintenance

Once the solution has been implemented, this section is designed to give you an understanding of the required upkeep. Reliability is another important factor in product adoption. Improvements/enhancements to the proposed solution should not result in unexpected downtime, and the solution should not require a high level of manual support to ensure proper operation.

- 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?
- Do you guarantee availability and uptime of the proposed solution to be 99.5%? How do you track system uptime?



Sustainability

- Can you provide an overview of your company's sustainability mission and strategy?
- How does sustainability align with your company's core values and business objectives?
- What specific sustainability goals has your company set, and how do you measure progress towards these goals?
- In what areas does your product have a major environmental impact?

Future Development

Ensure that the potential solution provider's roadmap is aligned with your organization's goals. Seeing the top priorities of technology development will provide clarity to the product direction and how your organization will be able to grow with the Industrial DataOps software.

- Can you provide a short-term (six to 12 months) and long-term (two to five years) product roadmap?
- What is your approach to developing new products and the possibilities for developing customizations/extensions?

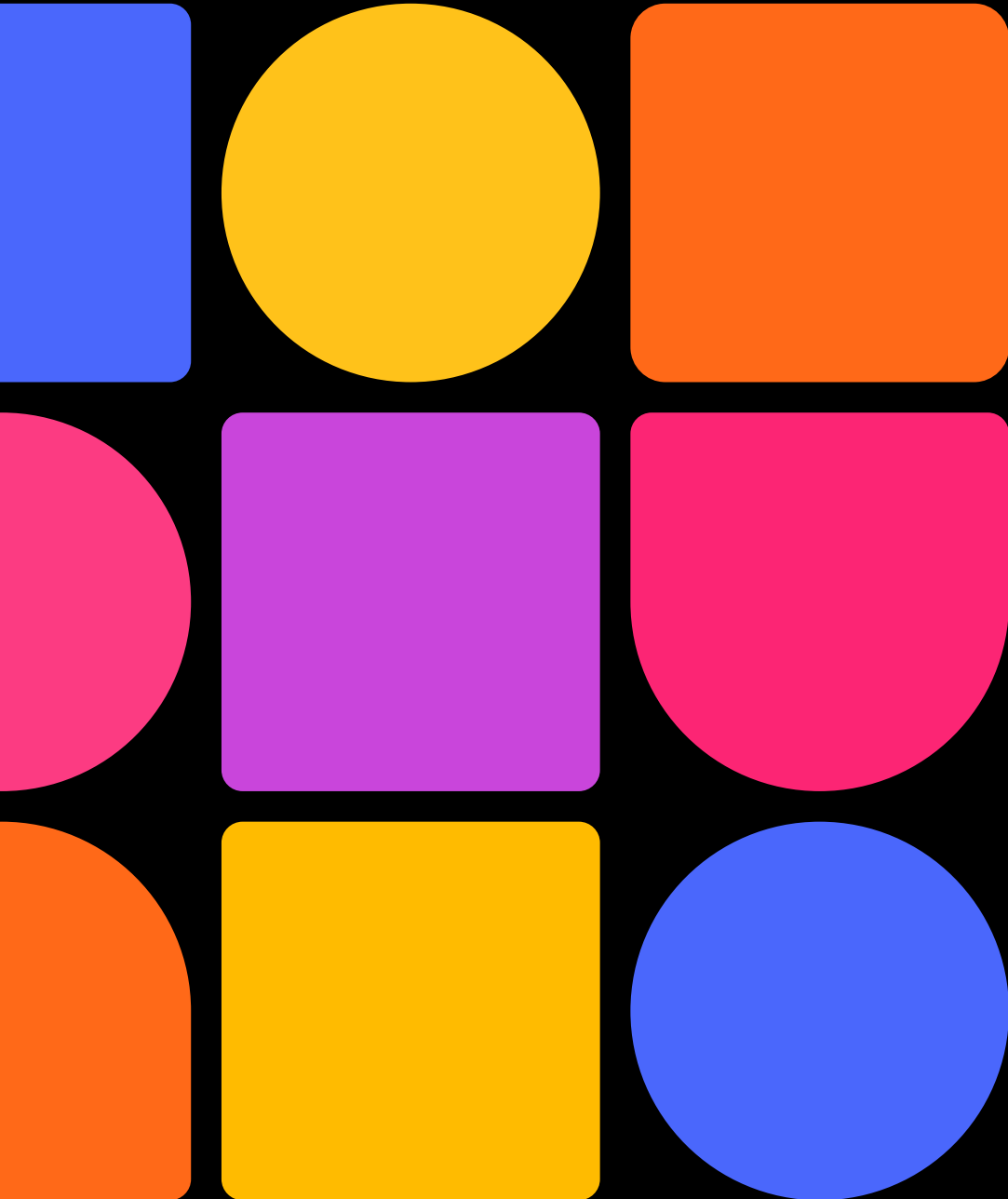


Pricing Model

To date, pricing has not seen convergence across 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.

- 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?





Learn more at cognite.ai →