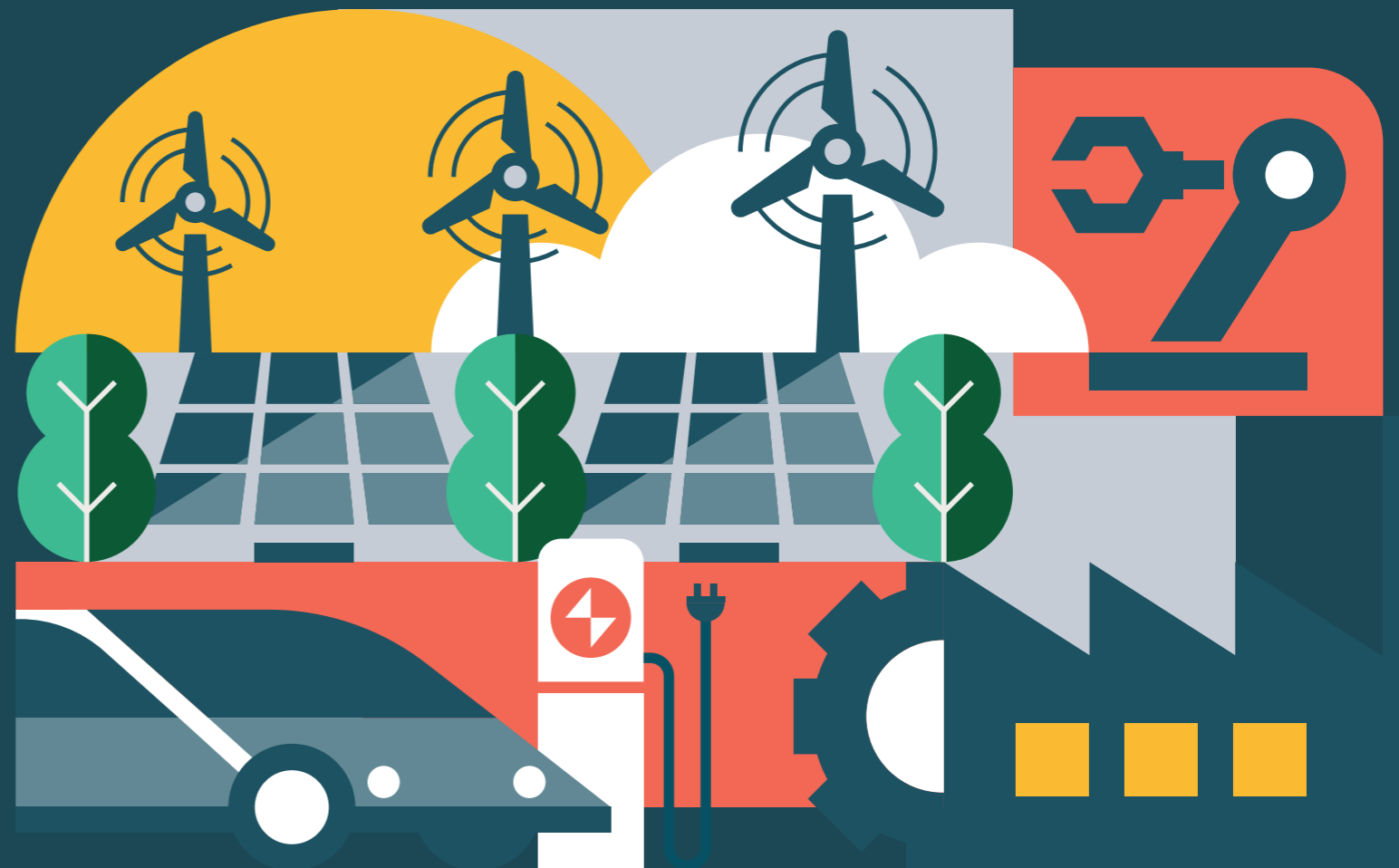


EBOOK

# Lakehouse for Manufacturing

Build a connected customer experience, optimize operations and unify your data ecosystem



# Contents



- Introduction ..... 3
- Manufacturing Transformation Trends ..... 5
- Manufacturing Data Challenges ..... 9
- Databricks Lakehouse for Manufacturing ..... 10
- Building Innovative Solutions on the Lakehouse ..... 12
  - SOLUTION: Part-Level Demand Forecasting ..... 12
  - SOLUTION: Overall Equipment Effectiveness & KPI Monitoring ..... 14
  - SOLUTION: Digital Twins ..... 15
  - SOLUTION: Computer Vision ..... 16
- An Ecosystem on the Lakehouse for Manufacturing ..... 17
  - SOLUTION: Avanade Intelligent Manufacturing ..... 18
  - SOLUTION: DataSentics Quality Inspector ..... 18
  - SOLUTION: Tredence Predictive Supply Risk Management ..... 19
- Leading Manufacturing Companies That Choose Us ..... 20

# Introduction



Market conditions in manufacturing are more challenging than ever. Operating margins and growth are impacted by the rising cost of labor, materials, energy and transportation, all peaking at the same time. Disruptive events in the supply chain are increasing in frequency and intensity, leading to significant revenue losses and damaged brand reputation.

Effective acquisition and retention of next-generation talent is a considerable issue for manufacturers. There are more jobs in the industry than there are people to do them, further compounding the problem of slower than expected industrial productivity growth over the last 15 years. The industry is also one of the largest consumers of energy, and faces a direct challenge of transforming operations to be more sustainable as governments are prioritizing net-zero policies that require a step change in energy efficiency and transition to low-carbon energy sources.

The manufacturing industry generates massive amounts of new data every day — estimated to be two to four times more in size than in industries such as communications, media, retail and financial services. This explosion of data has opened the door for the global manufacturing ecosystem to boost productivity, quality, sustainability and growth beyond what was previously thought possible.

Unfortunately, legacy data warehouse-based architectures weren't built for the massive volumes and type of data coming in through today's factories, products, processes and workers, let alone to support the advanced AI/ML use cases required to meet the customer expectations of shorter lead times, reliable delivery and smarter products.



For that, companies need to adopt a modern data architecture that provides the speed, scale and collaboration needed by broad teams of data engineers, data scientists, and analysts. Manufacturers need a comprehensive data platform that can not only handle massive volumes of data, but effectively and seamlessly operationalize the value from data, analytics and AI.

This is achieved by:

- Removing data silos by placing all data, regardless of type or frequency, in a single, open architecture — including unstructured data from sensors, telemetry, natural language logs, videos and images — helping you to gain end-to-end visibility into your business
- Ensuring your data is “always on” so that the freshest and highest quality data is available for all for the full spectrum of enterprise analytics and AI/ML use cases, allowing you to drive IT-OT convergence
- Having a comprehensive open architecture so IT and data teams can move with agility to bring AI and ML to where it’s needed, when it’s needed, including in connectivity-constrained environments
- Maintaining fine-grained governance and access control on your data assets, protecting sensitive intellectual property and customer data

The Databricks Lakehouse for Manufacturing does just this. It’s a comprehensive approach that empowers teams in the industry to collaborate and innovate around data, analytics and AI. It eliminates the technical limitations of legacy technologies and gives data teams the ability to drive deeper, end-to-end insight into supply chains, automate processes to reduce costs and grow productivity, and achieve sustainable transformation for a more prosperous future. Welcome to the Lakehouse for Manufacturing.

# Manufacturing Transformation Trends

The future of manufacturing is smart, sustainable and service oriented. Today's forward-thinking leaders are preparing the foundation they need to support that future by leveraging fast and connected data from all corners of the enterprise. There are four key trends driving transformation in manufacturing:

## Boosting industrial productivity through automation

A spike in labor costs, as well as the cost of energy and materials, puts significant pressure on operating margins. At the same time, industrial productivity has plateaued — it is at the same level today as it was in the late 2000s. In the face of these macro challenges and economic uncertainty, there has never been a more burning need to reduce costs and improve productivity through greater visibility and automation.

The industry has made strides in collecting data from machines and performing predictive analytics on sensor readings, with 47% of manufacturers citing the use of predictive maintenance to reduce operational costs with considerable upside ahead.

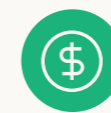
However, there is an entirely different class of unstructured data in the form of images, videos and LiDAR that is opening the door to game-changing automation in quality inspection, flow optimization and production scheduling. Historically, these critical processes have depended on manual and visual inspection of products and operations, which is resource intensive and less accurate than ML-driven computer vision techniques. This untapped data and capability is allowing manufacturers to deliver higher product quality and deliver on production demands using fewer resources. Andrew Ng, a machine learning

pioneer, rightly describes the massive opportunity for these technologies in his quote: "It is incumbent on every CEO in any manufacturing or industrial automation company to figure out how to make deep learning technology work for your business."

CUSTOMER STORY SPOTLIGHT:

Corning

”



**\$2 million** in cost avoidance through manufacturing upset event reduction

## Driving Better Efficiency in Manufacturing Process With ML

Corning has been one of the world's leading innovators in materials science for nearly 200 years. Delivering high-quality products is a key objective across the company's manufacturing facilities around the world, and it's always on a mission to explore how ML can help deliver on that goal. Databricks has been central to the company's digital transformation, as it provides a simplified and unified platform where teams can centralize all data and ML work. Now, they can train models, register them in MLflow, generate all additional artifacts — like exported formats — and track them in the same place as the base model.

LEARN MORE

## Gaining end-to-end operations and supply chain visibility

Modern customer expectations are forcing manufacturers to focus on more customer-centric KPIs: quality, on-time commitments and speed of delivery. That's not to say that asset and labor efficiency are less important — however, with customer expectations of shorter lead times and more reliable delivery, the success measures in manufacturing are shifting to a mantra of “measure what your customer values.”

High-performing manufacturers that embed this deep into their operational playbook also perform best on productivity and ROIC growth results, as evidenced in a recent study by the World Economic Forum and the International Centre of Industrial Transformation. The problem? In a post-pandemic world, operations and supply chains are persistently constrained, with increasing disruptions, spiraling costs and unpredictable performance. The business impact is considerable — studies have shown that a 30-day disruption can reduce EBITDA by 5% and impact annual revenue by as much as 20%.

Manufacturing companies need to be able to deliver on customer expectations, commitments and service levels, all while lowering costs and increasing productivity. Manufacturers need an enterprise data platform that can provide real-time visibility into order flows, production processes, supplier performance, inventory and logistics execution, breaking down departmental silos to maximize customer responsiveness, improve manufacturing agility and boost performance.

## Transforming your business model through tech-fueled services

Servitization, defined as the process of building revenue streams from services, has been trending for some time. The adaptation of the business model has been considerably profitable: on average, services account for ~30% of industrial manufacturing companies but contribute 60%+ of profit.

In aftersale services, a clear customer preference for business outcome-based offerings has emerged in almost every corner of the manufacturing industry. The use of data, analytics and AI is foundational to delivering more personalized customer outcomes, proactive field service delivery and differentiated mission-critical applications to their customers.

With greater autonomy, connectivity and sensorization, manufacturers operate in a paradigm where their products generate more and more data every second, opening up numerous new addressable opportunities for value creation. The business of manufacturing is no longer linear, and manufacturers will need to reimagine their businesses to go beyond merely providing the primary unit of production — the next SKU, machine, vehicle or airplane — and leverage this data to operate a platform business with higher growth, stickier revenue streams and greater resilience to demand shocks.

CUSTOMER STORY SPOTLIGHT:

## Rolls-Royce

”

### Aerospace Goes Green With Data and AI

While most people think of luxury cars when they hear “Rolls-Royce,” the Civil Aerospace branch is its own company, having separated from the car manufacturing arm in 1971. The now wildly successful manufacturer of commercial airplane engines is a leader in its industry for innovation. Today, Rolls-Royce obtains information directly from the airlines’ engines and funnels it into the Databricks platform. This gives the company insights into how the engines are performing and ways to improve maintenance schedules, translating to less downtime, delays, and rerouting — all of which reduce carbon footprint.

[LEARN MORE](#)

*“We employed Databricks to optimize inventory planning using data and analytics, positioning parts where they need to be, based on the insight we gain from our connected engines in real time and usage patterns we see in our service network. This has helped us minimize risks to engine availability, reduce lead times for spare parts and drive more efficiency in stock turns — all of this enables us to deliver TotalCare, the aviation industry’s leading Power-by-the-Hour (PBH) maintenance program.”*

#### STUART HUGHES

Chief Information and Digital Officer  
Rolls-Royce Civil Aerospace



## Driving a more sustainable approach to manufacturing

Global efforts on reducing greenhouse gas (GHG) emissions are accelerating, with over 70 countries representing more than 75% of global emissions having signed agreements to reach net-zero emissions by 2050. Manufacturing-centric sectors are critical to achieving net-zero sustainability commitments around the world, as they represent over 50% of global energy consumption and contribute to ~25% of global emissions.

Those at the forefront of data, analytics and AI are setting science-based targets and are driving favorable sustainability outcomes today by deriving better insights from their operations, supply chains and the outcomes that their products generate for their end customers.

### CUSTOMER STORY SPOTLIGHT:

## Shell



### Delivering Innovative Energy Solutions for a Cleaner World

Shell has been at the forefront of creating a cleaner tomorrow by investing in digital technologies to tackle climate change and become a net-zero emissions energy business. Across the business, they are turning to data and AI to improve operational efficiencies, drive customer engagement, and tap into new innovations like renewable energy. Hampered by large volumes of data, Shell chose Databricks to be one of the foundational components of its Shell.ai platform. Today, Databricks empowers hundreds of Shell's engineers, scientists and analysts to innovate together as part of their ambition to deliver cleaner energy solutions more rapidly and efficiently.

[LEARN MORE](#)

*"Shell has been undergoing a digital transformation as part of our ambition to deliver more and cleaner energy solutions. As part of this, we have been investing heavily in our data lake architecture. Our ambition has been to enable our data teams to rapidly query our massive data sets in the simplest possible way. The ability to execute rapid queries on petabyte scale data sets using standard BI tools is a game changer for us. Our co-innovation approach with Databricks has allowed us to influence the product road map, and we are excited to see this come to market."*

#### DANIEL JEAVONS

General Manager – Advanced Analytics CoE  
Shell

**Millions**  
of dollars saved in  
potential engine  
repair costs

**250** data team  
members supporting  
160+ high-value use  
cases

**9x** faster –  
5 minutes to validate  
a label, reduced from  
45 minutes



# Manufacturing Data Challenges

## ➤ Massive unstructured/OT data volumes

The industry is seeing immense growth in data volumes: much of this massive growth is due to semi-structured and unstructured data from connected workers, buildings, vehicles and factories. This growth in multi-modal data from IoT sensors, process historians, product telemetry, images, cameras and perception systems has outpaced legacy data warehouse-centric technologies. On-prem and cloud data warehouse tech-based architectures are too complex and too costly for the large and heterogeneous data sets prevalent in the industry.

## ➤ Driving IT-OT convergence

The success and pace of data modernization efforts in manufacturing is so often muted by critical data being stuck in multiple closed systems and proprietary formats, making it difficult and cost-prohibitive to extract the full potential of IT and OT data sets. In addition, data quality issues such as outdated or inaccurate data can often lead to a disjointed and incomplete view of customers, operations and assets. For years, companies have lacked a common foundation for complex and heterogeneous manufacturing data — from IoT-generated data streams to financial metrics stored in ERP applications — and it has impacted their ability to provide the freshest, highest-quality and most complete data for analytics.

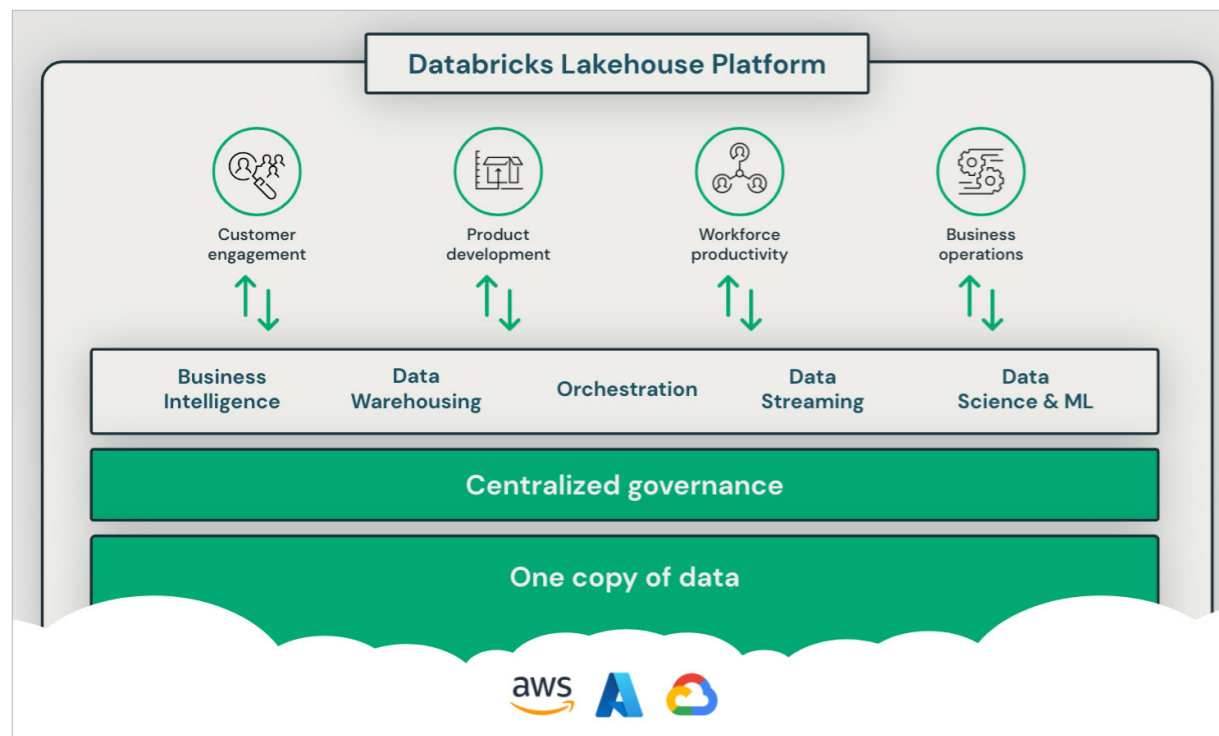
## ➤ Bringing AI/ML to where it's needed

To realize the promise of AI/ML in manufacturing, machine learning models need to be brought as close to the decision as possible, often at the edge in facilities and locations with limited or intermittent connectivity to the internet or cloud. This requires deployment flexibility to on-premises or edge devices, with an experience comparable to that in the cloud.

## ➤ Inability to innovate at scale

CDOs want to be able to quickly and efficiently reproduce successes at global scale. Technical and business users want to simply and quickly know what data sets are available to solve the business issue at hand. Analysts want flexibility to use the tools they are most familiar with in order to stay responsive to business needs. Fragmented approaches to architecture and tooling make scaling business impact very difficult, which results in talent churn, slower development and duplicative efforts — all leading to higher costs.

# Databricks Lakehouse for Manufacturing



## ▶ Deliver personalized outcomes and frictionless experiences

- Millions of assets streaming IoT data
- 5%–10% reduction in unplanned downtime and cost
- Accurate prices across 1,000s of locations and millions of dealers
- 200%+ increase in offer conversion rates

With Databricks Lakehouse for Manufacturing, manufacturers can gain a single view of their customers that combines data from each stage of the customer journey. With a 360-degree view in place, manufacturers can drive more differentiated sales strategies and precise service outcomes in the field, delivering higher revenue growth, profitability and CSAT scores.

With the Databricks Lakehouse, you can analyze product telemetry data, customer insights and service networks to deliver highest uptime, quality of service and economic value through the product lifecycle.

## Optimize the supply chain, production processes and fulfillment logistics with real-time analytics and AI.

The Databricks Lakehouse for Manufacturing is the only enterprise data platform that helps manufacturing organizations optimize their supply chains, boost product innovation, increase operational efficiencies, predict fulfillment needs and reduce overall costs.

### Gain real-time insight for agile manufacturing and logistics

- 30%–50% improvement in forecast accuracy
- 90% lower cost for new manufacturing line
- 4%–8% reduction in logistics costs
- 10% improvement in carbon footprint

The Databricks Lakehouse lets you build a resilient and predictive supply chain by eliminating the trade-off between accuracy or depth of analysis and time. With scalable, fine-grained forecasts to predict or sense demand, or perform supply chain planning and optimization, Databricks improves accuracy of decisions, leading to higher revenue growth and lower costs.

The lakehouse provides an “always on” architecture that makes IT-OT convergence a reality, by continuously putting all data to work regardless of the frequency at which it arrives (periodic, event-driven or real-time streaming) and creates valuable data products that can empower decision makers. This creates real-time insight into performance with data from connected factory equipment, order flows and production processes to drive the most effective resource scheduling.

### Empower the manufacturing workforce of the future

- 25% improvement in data team productivity
- 50x faster time to insight
- 50% reduction in workplace injuries

With Databricks, manufacturers can increase the impact and decrease the time-to-value of their data assets, ultimately making data and AI central to every part of their operation. And by empowering data teams across engineering, analytics and AI to work together, Databricks frees up employees to self-serve and focus on realizing maximum business value — improving product quality, reducing downtime and exceeding customer expectations.

### Execute product innovation at the speed of data

- 90% decrease in time to market of new innovations
- 20x faster data processing of vehicle and road data

It is critical that manufacturers are offering the most desirable value propositions so end consumers don't look elsewhere. By tapping into product performance and attribute data along with market trends and operations information, manufacturers can make strategic decisions.

With Databricks, manufacturers can decrease time to market with new products to increase sales by analyzing customer behavior and insights (structured, unstructured and semi-structured), product telemetry (streaming, RFID, computer vision) and digital twins, and leveraging that data to drive product decisions.

# Building Innovative Solutions on the Lakehouse

The flexibility of the Databricks Lakehouse Platform means that you can start with the use case that will have the most impact on your business. Through our experience working with some of the largest and most cutting-edge manufacturers in the world, we've developed Solution Accelerators based on the most common needs of manufacturers to help you get started. These purpose-built guides — fully functional notebooks and best practices — speed up results across your most common and high-impact use cases. Go from idea to proof of concept (PoC) in as little as two weeks. Check out the full list of Solution Accelerators [here](#).

## SOLUTION Part-Level Demand Forecasting



Demand forecasting is a critical business process for manufacturing and supply chains. McKinsey estimates that over the next 10 years, supply chain disruptions can cost close to half (~45%) of a year's worth of profits for companies. Having accurate and up-to-date forecasts is vital to plan the scaling of manufacturing operations, ensure sufficient inventory and guarantee customer fulfillment.

In recent years, manufacturers have been investing heavily in quantitative-based forecasting that is driven by historical data and powered using either statistical or machine learning techniques. Benefits include:

- Better sales planning and revenue forecasting
- Optimized safety stock to maximize turn-rates and service-delivery performance
- Improved production planning by tracing back production outputs to raw material levels

➤ **A disruption lasting just 30 days or less could equal losses of 3%–5% of EBITDA.**

Databricks Lakehouse can enable large-scale forecasting solutions to help manufacturers navigate the most common data challenges when trying to forecast demand.

#### COMMON USE CASES:

- Scalable, accurate forecasts across large numbers of store-item combinations experiencing intermittent demand
- Automated model selection to ensure the best model is selected for each store-item combination
- Metrics to identify the optimal frequency with which to generate new predictions
- Manage material shortages and predict overplanning

Try our **Parts-Level Solution Accelerator** to facilitate fine-grained demand forecasts and planning.



## SOLUTION

Overall Equipment Effectiveness  
& KPI Monitoring

The need to monitor and measure manufacturing equipment performance is critical for operational teams within manufacturing. Today, Overall Equipment Effectiveness (OEE) is considered the standard for measuring manufacturing equipment productivity. According to Engineering USA, an OEE value of 85% or above is considered world-leading. However, many manufacturers typically achieve a range of between 40% and 60%. Reasons for underachievement often include:

- Delayed inputs due to manual processes that are prone to human error
- Bottlenecks created by data silos, impeding the flow of fresh data to stakeholders
- A lack of collaboration capabilities, keeping stakeholders from working on the same information at the same time

➤ **Poor OEE value can be a result of poor parts quality, slow production performance and production availability issues.**

Databricks Lakehouse can help manufacturers maneuver through the challenges of ingesting and converging operational technology (OT) data with traditional data from IT systems to build forecasting solutions.

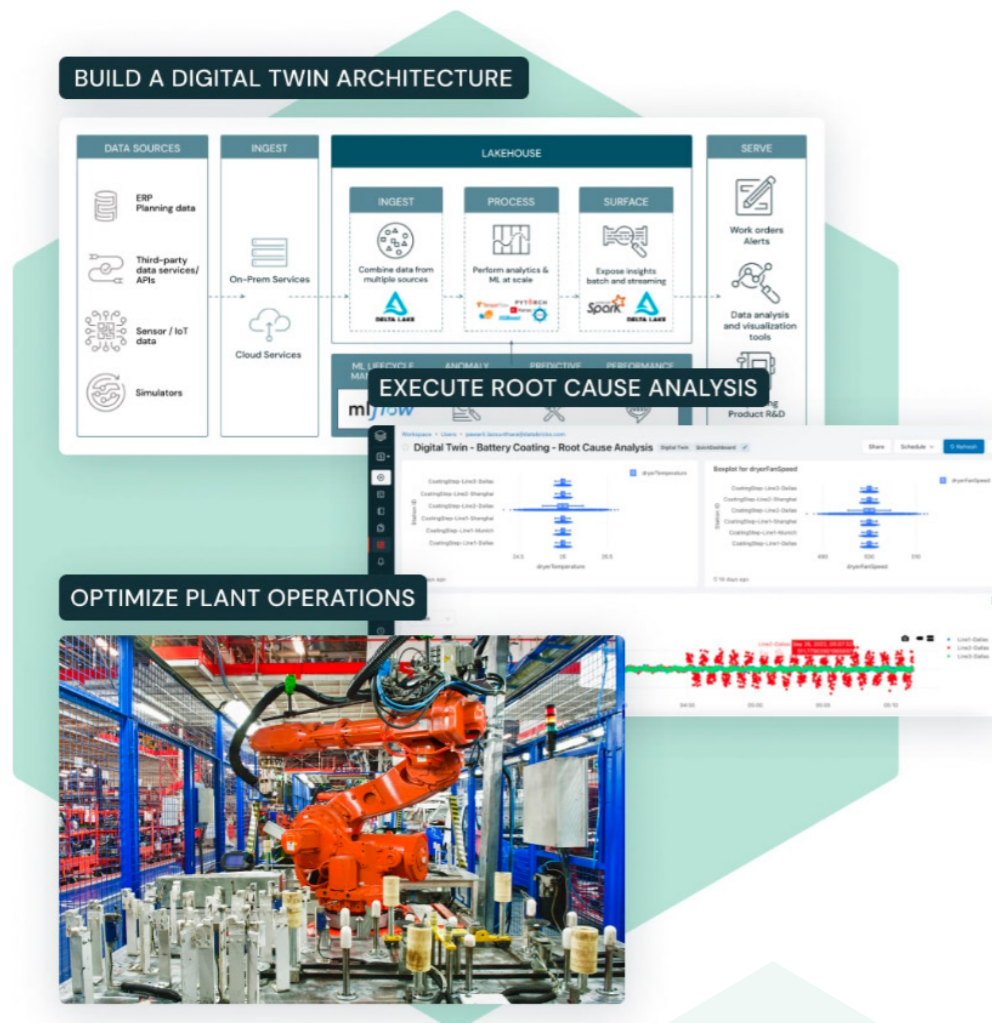
## COMMON USE CASES

- Incrementally ingest and process sensor data from IoT devices in a variety of formats
- Compute and surface KPIs and metrics to drive valuable insights
- Optimize plant operations with data-driven decisions

Try our **Solution Accelerator for OEE and KPI Monitoring** for performant and scalable end-to-end monitoring.

## SOLUTION

### Digital Twins



Market dynamics and volatility are requiring manufacturers to bring products to market more quickly, optimize production processes and build agile supply chains at scale at a lower price. To do so, many manufacturers have turned to building digital twins, which are virtual representations of objects, products, pieces of equipment, people, processes or even complete manufacturing ecosystems.

Digital twins provide insights — derived from sensors (often IoT or IIoT) that are embedded in the original equipment — that have the potential to transform the manufacturing industry by driving greater efficiency, reducing costs and improving quality.

➤ **Digital twin technologies can improve product quality by up to 25%.**

Databricks Lakehouse can bring digital twins to life through fault-tolerant processing of streaming workloads generated by IoT sensor data and complex event processing (important for modeling physical processes).

#### COMMON USE CASES

- Process real-world data in real time
- Compute insights at scale and deliver to multiple downstream applications
- Optimize plant operations with data-driven decisions

Try our **Solution Accelerator for Digital Twins** to accelerate time to market of new innovations.

**SOLUTION****Computer Vision**

The rise in computer vision has been fueled by the rapid developments in neural network technologies, which use AI to better understand and interpret images with near-perfect precision. In manufacturing, computer vision can transform operations by, for example, identifying product defects to improve quality control, detecting safety hazards on the production floor, and tracking and managing inventory levels.

 **As per the American Society for Quality, cost of poor quality for companies can be as high as 20% of revenue.**

Databricks Lakehouse can easily ingest complex, unstructured image and video data at massive scale. Through the most popular computer vision libraries, data teams can scale AI models that leverage computer vision to recognize patterns, detect objects and make predictions with 99% accuracy.

**COMMON USE CASES**

- Quickly identify defects and ensure that products and processes meet quality standards
- Automate positioning and guidance to ensure that parts and products are properly aligned and assembled
- Predict maintenance issues to reduce downtime and maintenance costs, improve parts reliability, and increase safety for workers

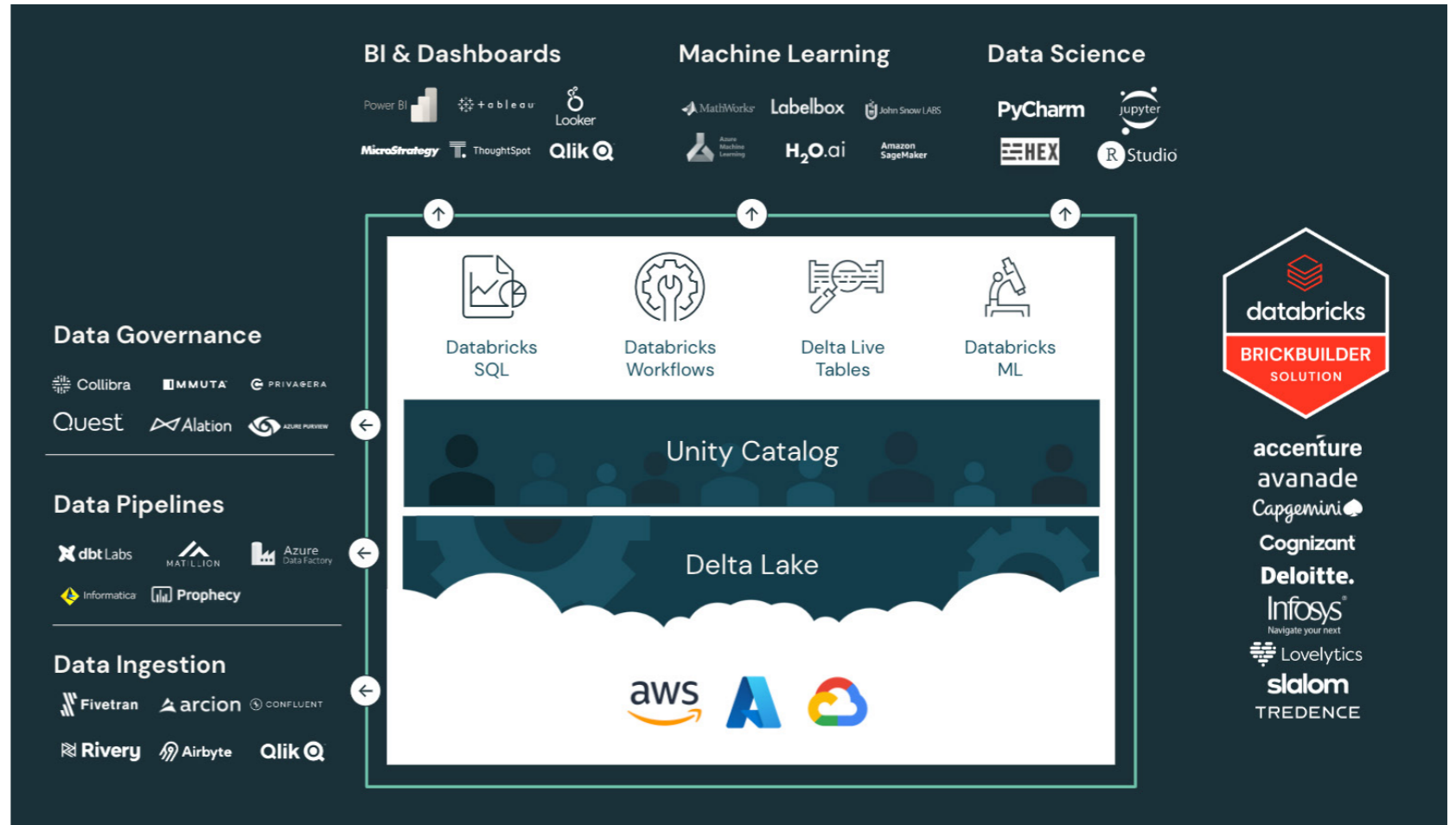
Try our **Solution Accelerator for Computer Vision** to improve efficiency, reduce costs and enhance overall safety.



# An Ecosystem on the Lakehouse for Manufacturing

We've partnered with leading consulting firms and independent software vendors to deliver innovative, manufacturing-specific solutions. Databricks Brickbuilder Solutions help you cut costs and increase value from your data. Backed by decades of industry expertise — and built for the Databricks Lakehouse Platform — Brickbuilder Solutions are tailored to your exact needs.

We also work with technology partners like Alteryx, AtScale, Fivetran, Microsoft Power BI, Qlik, Sigma, Simplement, Tableau and ThoughtSpot to accelerate the availability and value of data. This allows businesses to unify data from complex source systems and operationalize it for analytics, AI and ML on the Databricks Lakehouse Platform.



**SOLUTION****Avanade Intelligent Manufacturing**

Every year, businesses lose millions of dollars due to equipment failure, unscheduled downtime and lack of control in maintenance scheduling. Along with lost dollars, businesses will experience lower employee morale when stations are in and out of service. Avanade's Intelligent Manufacturing solution supports connected production facilities and assets, workers, products and consumers to create value through enhanced insights and improved outcomes. Manufacturers can harness data to drive interoperability and enhanced insights at scale using analytics and AI. Outcomes include improvements across production (e.g., uptime, quantity and yield), better experiences for workers, and greater insight into what customers want.

Try our joint solution, **Intelligent Manufacturing**, to drive value and operationalize team coordination and productivity.

**SOLUTION****DataSentic Quality Inspector**

Quality control is a crucial aspect of any production process, but traditional methods can be time-consuming and prone to human error. Quality Inspector by DataSentic, an Atos company, offers a solution that is both efficient and reliable. With out-of-the-box models for visual quality inspection, which are tailored to meet specific business requirements, organizations will experience stable, scalable quality control that's easy to improve over time. Quality Inspector is an end-to-end solution that can be seamlessly integrated into an existing setup, delivering high performance and reliability.

Try our joint solution, **Quality Inspector**, to automate production quality control with an increase in accuracy and quicker time to value.

## SOLUTION

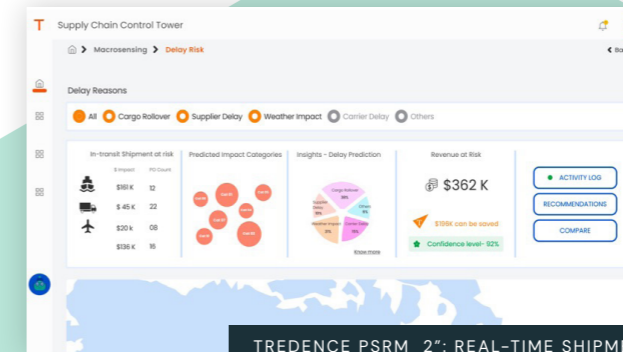
## Tredence Predictive Supply Risk Management

Customers today are faced with multiple supply risks including lack of in-transit visibility, disruptions caused by weather, local events, among others. Tredence's Predictive Supply Risk Management solution, built on the Databricks Lakehouse Platform, helps businesses meet supply risk challenges by providing a scalable, cloud-based solution that can be tailored to the specific needs of each organization. The platform's flexibility and scalability allow businesses to keep pace with changing regulations and customer demands, while their comprehensive suite of tools helps identify and mitigate risks across the enterprise.

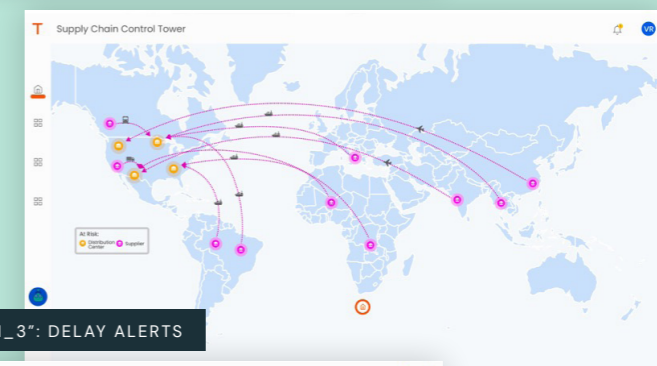
Try our joint solution, **Predictive Supply Risk Management**, to predict order delays, identify root causes and quantify supply chain impact.

Visit our [site](#) to learn more about our Databricks Partner Solutions.

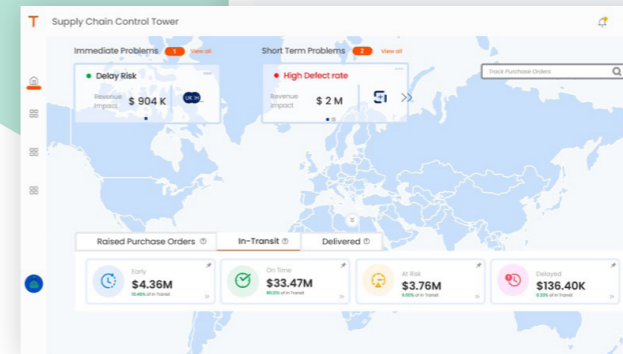
## TREDENCE PSRM\_1": PREDICT SUPPLY RISK



## TREDENCE PSRM\_2": REAL-TIME SHIPMENT VISIBILITY



## TREDENCE PSRM\_3": DELAY ALERTS



# Leading Manufacturing Companies That Choose Us

 RIVIAN	 TOYOTA	 BRIDGESTONE	 CISCO	 Cummins	 DELL	 Electrolux	 hp
 Husqvarna	 J.B. HUNT	 IVECO	 Mercedes-Benz	 Raytheon Technologies	 Rockwell Automation	 ROLLS ROYCE	 SIEMENS
 STELLANTIS	 ThermoFisher SCIENTIFIC	 VIESSMANN	 Intelematics	 Miele	 Royal IHC	 AGC	 ARTC
	 LENNOX	 MARELLI	 mipox	 oldendorff EO	 Virgin hyperloop one	 incite	

Databricks is the lakehouse company. More than 9,000 organizations worldwide — including Comcast, Condé Nast and over 50% of the Fortune 500 — rely on the Databricks Lakehouse Platform to unify their data, analytics and AI. Databricks is headquartered in San Francisco, with offices around the globe. Founded by the original creators of Apache Spark,™ Delta Lake and MLflow, Databricks is on a mission to help data teams solve the world’s toughest problems. To learn more, follow Databricks on [Twitter](#), [LinkedIn](#) and [Facebook](#).

Get started with a free trial of Databricks and start building data applications today

START YOUR FREE TRIAL

To learn more, visit us at:

[Manufacturing Industry Solutions](#)

