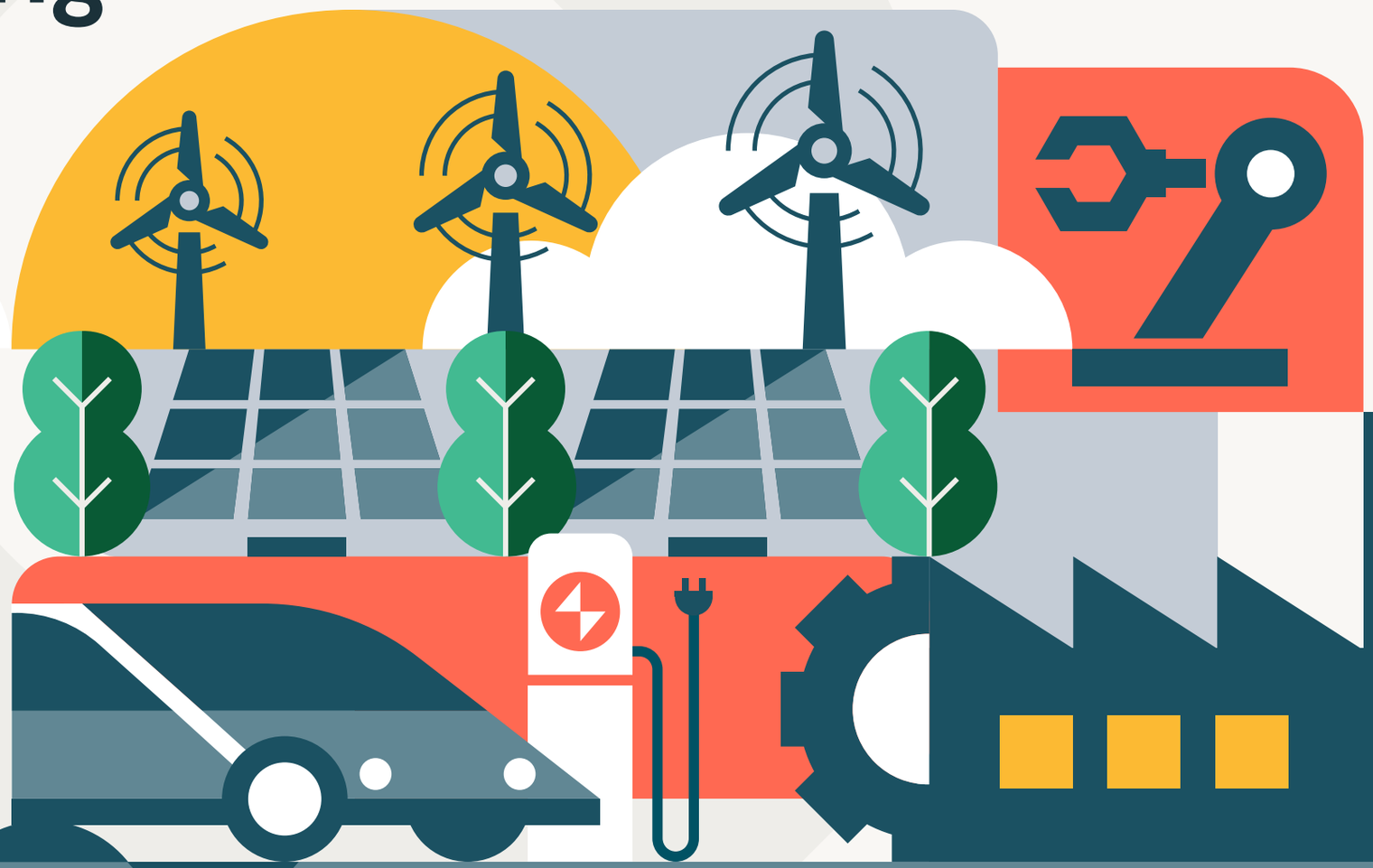
eBook

From Bits to Bolts:

Navigating the IoT Data Landscape in Manufacturing



databricks

Contents

CHAPTER 1: **Introduction** 3

CHAPTER 2: **Applying IoT Across Manufacturing Use Cases** 4

Innovating and accelerating product design 4

Powering positive aftermarket service experiences 4

Enhancing asset performance monitoring 5

Gaining visibility into smart manufacturing 6

CHAPTER 3: **Challenges in the IoT Data Landscape** 7

Large volumes of complex, unstructured data 7

Data curation across disparate systems 7

High-velocity real-time data streaming 8

Data security and governance 9

CHAPTER 4: **Lakehouse for Manufacturing Simplifies and Accelerates IoT Use Cases Across the Value Chain** 10

Solution Accelerator: Multi-factory OEE and KPI monitoring done fast 10

Solution Accelerator: Digital twins enable the safe testing of new ideas 11

Solution Accelerator: Scalable route generation drives profitable delivery 12

IIoT technologies fuel prediction with precision 13

CHAPTER 5: **Next Steps Toward IoT-Supported Manufacturing** 14

CHAPTER 1:

Introduction

Manufacturing has evolved significantly over the years, and today, with the emergence of innovations, it stands on the brink of a new era. This era is introducing novel approaches to streamline operations and accelerate time to market. The key to fueling these new innovations is data, most of which can be generated by the Internet of Things (IoT), which integrates various interconnected objects, sensors, machines and systems within the industrial environment. Enterprises can enhance supply chain efficiency, accelerate productivity and make more informed decisions by harnessing the power of unstructured, high-velocity sensor data.

Of course, these new capabilities are a major catalyst for wins across industries like healthcare and life sciences, transportation, and retail and consumer goods. Now, manufacturers can monitor performance, track production variables, predict maintenance needs, and optimize resource allocation by seamlessly connecting machinery, equipment, and people to the internet, data platforms, and innovative tools. With these capabilities, enterprises can rapidly meet changing consumer preferences while streamlining operations, fulfilling orders faster and, ultimately, reducing overall costs.

While IoT isn't a new concept, it's quickly gaining adoption with accessible data platforms and AI tooling to extract actionable insights. Due to the explosion of connected devices and the deluge of unstructured data ingested daily, manufacturers have struggled with efficiently capturing, preparing and operationalizing data.

Enterprises that don't take notice of the competitive advantage of IoT and modernize data management accordingly risk falling behind in product quality, time to market, and customer satisfaction. It's not just a buzzword anymore. IoT is more relevant and realistic to organizational growth than ever before.

Gartner

"The IT Services for IoT market will represent a $58 billion opportunity in 2025, up at a 34% CAGR from 2020."

— Gartner Research →

CHAPTER 2:

Applying IoT Across Manufacturing Use Cases

Innovating and accelerating product design

IoT presents a powerful opportunity for organizations to advance product design with new product development insights for customization and speed. With the constant feedback loop between IoT-connected devices and the enterprise, engineers, data scientists and analysts can leverage product performance in real-world scenarios for iterative design improvements.

The data captured by IoT devices allows for remote monitoring and diagnostics during manufacturing, ensuring product quality and equipment maintenance are serviced effectively. That enhances product longevity for customer satisfaction and extends the lifespan of expensive machinery. Taking advantage of IoT, enterprises can embark on a journey of agile design, rapid prototyping and continuous systems improvement. The result is products that align with customer needs and preferences while staying competitive in an ever-evolving market landscape.

Powering positive aftermarket service experiences

Enterprises must always aim to provide a flexible and forward-thinking experience that anticipates potential customer needs and wants. With IoT, manufacturers can gather continuous data on equipment health with sensors that monitor parameters like temperature, vibration and energy consumption. When deviations from normal operating conditions are detected, predictive analytics can anticipate potential breakdowns, allowing service teams to intervene before a major issue occurs. Not only does this minimize expensive downtime, but it also enhances your reputation for reliable products and responsive support.

Connected vehicles with IoT sensors transmit real-time data about engine performance, tire pressure, coolant temperature, driver behavior and overall vehicle health. Manufacturers can analyze this data to identify patterns and predict maintenance needs. For instance, if a vehicle's diagnostic data suggests that a specific part is nearing the end of its lifespan, the manufacturer can proactively schedule a service appointment, order the necessary replacement parts, and ensure that the customer's experience is smooth and hassle-free. Predictive maintenance preserves and extends the life and value of costly machinery by reducing the likelihood of a breakdown and optimizing the scheduling of service resources.

Furthermore, IoT-enabled products facilitate diagnostics and troubleshooting. Medical equipment manufacturers can remotely access diagnostic data from IoT-equipped devices within healthcare facilities. When an issue arises, technicians can often diagnose the problem remotely, saving time and resources that would otherwise be spent on travel and onsite services. This streamlined process accelerates issue resolution, minimizes equipment downtime and ensures critical healthcare systems remain operational — benefiting manufacturers and end users alike.

Enhancing asset performance monitoring

Data is a powerful asset for decision-making processes, and with IoT, organizations can gather, analyze and apply massive data sets generated throughout nearly every aspect of the production process. IoT-enabled devices can communicate and connect to surrounding devices and infrastructures to enhance quality control. By embedding sensors in various stages of production, manufacturers get continuous data on dimensions, tolerances and material properties. Manufacturers can implement automated quality checks, issue detection, and alerts using a data platform to stop potential defects early on and avoid waste.

IoT-enabled asset monitoring is also pivotal for compliance within highly regulated industries like consumer packaged goods (CPGs) and pharmaceuticals. Manufacturers can use sensors to monitor critical environmental conditions while producing and storing sensitive products. For example, in pharmaceutical manufacturing, IoT devices can track temperature, humidity and other variables in real time. Constant monitoring ensures that products are manufactured and stored within specified parameters to maintain efficacy and meet regulatory standards. Automated alerts can trigger immediate corrective actions when an anomaly is identified, preventing costly compliance violations and safeguarding product integrity.

Gaining visibility into smart manufacturing

IoT has been transformative in turning traditional factories into smart factories to improve various production processes. To monitor and optimize energy usage and help enterprises decrease their carbon footprint, manufacturers can deploy IoT sensors that track consumption patterns. With 5G, GPS, machine-to-machine (M2M) and machine-to-people (M2P) communications, factories enhance workplace safety by identifying hazards and alerting workers to potential dangers. Layering these insights with AI and ML, organizations can implement automatic self-remediation, alerting and escalation policies to prevent business disruptions.

Within supply chain management, IoT-equipped devices can track the movement and location of raw materials, components and finished products across processes. With visibility into the entire production journey, manufacturers can accurately predict demand, optimize inventory management and streamline logistics planning. They can also respond quickly to changes in demand, manage supply, allocate resources effectively and make on-time deliveries to satisfy customers while minimizing costs.

Feeding data and analytics through a data platform allows organizations to visualize different aspects of manufacturing with real-time dashboards that provide an immediate overview of key performance indicators (KPIs). With these insights, manufacturers can identify bottlenecks, monitor progress and make data-driven decisions to improve process efficiencies. For example, in an automotive manufacturing plant, IoT sensors can monitor and analyze various parameters across the assembly line to ensure consistent quality, minimize defects and manage labor through real-time process adjustments.

CHAPTER 3:

Challenges in the IoT Data Landscape

"By 2025, the total data volume of connected IoT devices worldwide is forecast to reach 79.4 zettabytes (ZBs)."

— Statista →

Large volumes of complex, unstructured data

As sensors, devices and interconnected systems generate copious amounts of real-time data, the amount can overwhelm existing data management infrastructure, causing it to buckle under the weight of expectations. Larger enterprises can process over a billion data elements daily from more than a million connected assets worldwide, and there's no sign of slowing down. IoT data is expected to grow exponentially as sensors become more complex, and organizations must scale horizontally to keep up.

In a smart factory, for example, each sensor on each machine generates various data sets on performance, temperature, vibration, security and operational status. This continuous influx of data can strain storage and processing capabilities on traditional processing solutions. Moreover, the complexity of this data, often arriving in different formats and from diverse sources, makes integration and analysis a formidable task. In a supply chain context, IoT-connected devices track the movement of goods through complex data streams that need to be synchronized and interpreted for inventory management and deliveries. As organizations continue to balance the need for immediate insights with the ability to manage and process large volumes of diverse data, enterprises must modernize their data infrastructures.

Data curation across disparate systems

Combining sensor and IoT data with other data sources — including enterprise data stored in data warehouses, industrial control systems and business applications — is essential for realizing IoT use cases. Unfortunately, data, data teams and business users are historically siloed within disparate systems, compromising data completeness and integrity. Data teams then have to build an entirely new tech stack — stitching together different services tailored to each use case — slowing productivity, time to insights, and innovation. Integrating heterogeneous data sources requires high-quality data to avoid incorrect insights that can negatively impact decision-making and damage business performance.

In manufacturing, IoT sensors might generate data that includes both critical operational insights and irrelevant noise. Curating this data to filter out extraneous information while preserving vital details necessitates robust data cleansing and preprocessing methods. Similarly, integrating data from vendors, locations and systems in a connected supply chain requires standardized formats and consistent labeling for coherent analysis. In healthcare settings, predictive algorithms for suicide prevention and preventative medical screenings must quickly ingest and process millions of patient records across locations — including medical history data like missed mental health appointments, prescribed medications, and risk factors across hundreds of thousands of drugs and diseases. For these advanced analytics, enterprises must implement specialized and scalable tools and technologies that enable efficient and cost-effective data curation across sources.

High-velocity real-time data streaming

The speed of streaming IoT data can be challenging for enterprises to manage due to the rapid influx of information that must be captured, processed, analyzed and acted upon at the rate it's generated. In addition to streaming data, organizations also need low-latency data and model serving to power the real-time decisions necessary for short service-level agreements (SLAs). Immediate insights into operations, performance, supply and customer engagement are critical use cases for IoT data, but processing delays can damage end products, interrupt deliveries and lower customer satisfaction.

In applications like financial trading, where IoT sensors gather real-time market data to inform algorithmic decisions, even the slightest delay in data processing can lead to significant financial losses. In healthcare, wearable devices continuously monitor patient vital signs, generating a constant stream of data that demands immediate analysis for timely medical interventions. Furthermore, in logistics and transportation, IoT sensors track real-time vehicle movements, shipment status, traffic patterns, and driving behaviors to optimize routes, delivery schedules, driver safety, and vehicle performance. Any lags in data processing or accessing historical data could result in costly mistakes and unnecessary waste. The challenge lies in developing a robust data infrastructure capable of handling high-velocity data flows, implementing sophisticated algorithms, and ensuring seamless communication between devices and data systems.

Data security and governance

The intricate web of IoT interconnected devices and the diverse data they generate require strong security measures to safeguard data transmission, device authentication and data storage. Ensuring proper governance over data ownership, access, usage and compliance is also required, but it gets complicated quickly. In ecosystems where data moves between devices, platforms and organizations, enterprises struggle to maintain stringent security and governance standards efficiently — but it may be required.

In the airline industry, connected aircraft data needs the approval of pilot unions before being utilized. For automotive companies, vehicle identification numbers (VINs) are considered personally identifiable information (PII) in many countries and require additional protections and management. Shop floor data also contains sensitive intellectual property (IP) information critical to the business's competitive advantage. For enterprises to benefit from IoT, they must implement security and governance controls capable of protecting business-critical information regardless of where the data is.

CHAPTER 4:

Lakehouse for Manufacturing Simplifies and Accelerates IoT Use Cases Across the Value Chain

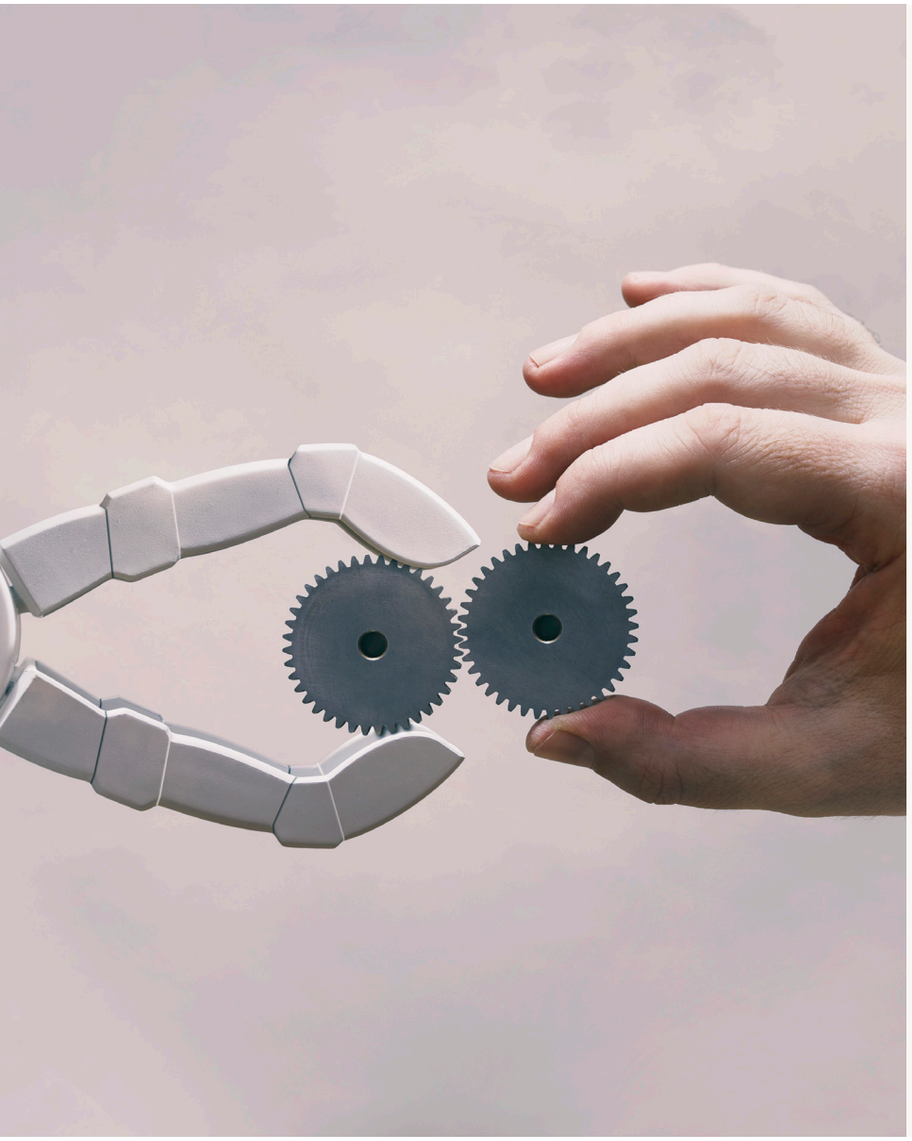
Leveraging Databricks Lakehouse for Manufacturing, enterprises can unify data and seamlessly integrate data engineering, data science and business analytics to efficiently manage the volume, velocity and variety of IoT data. With Databricks Solution Accelerators, data teams can significantly speed the process from idea to proof of concept using purpose-built guides with fully functional notebooks and best practices. See how these components come together to fulfill IoT use cases in manufacturing with precision, ease and end-to-end visibility.

SOLUTION ACCELERATOR

Multi-factory OEE and KPI monitoring done fast

Overall Equipment Effectiveness (OEE) is the standard for measuring manufacturing equipment productivity. Monitoring OEE and other KPIs has traditionally been a manual exercise, making it difficult to compute the latency and scale required within legacy systems. With Databricks' medallion architecture and the Databricks Solution Accelerator for OEE, enterprises can achieve performant, scalable and real-time equipment monitoring across manufacturing processes and within multi-factory environments.

- **Incrementally ingest and process data from sensors and IoT devices** in various formats for broader applications and more visibility across the organization
- **Compute and surface KPIs and metrics faster and more efficiently** to drive productivity, insights, optimizations and innovation with valuable insights
- **Improve plant operations with data-driven decisions** based on current availability, past performance and underperforming indicators, and perform root cause analysis for risk mitigation

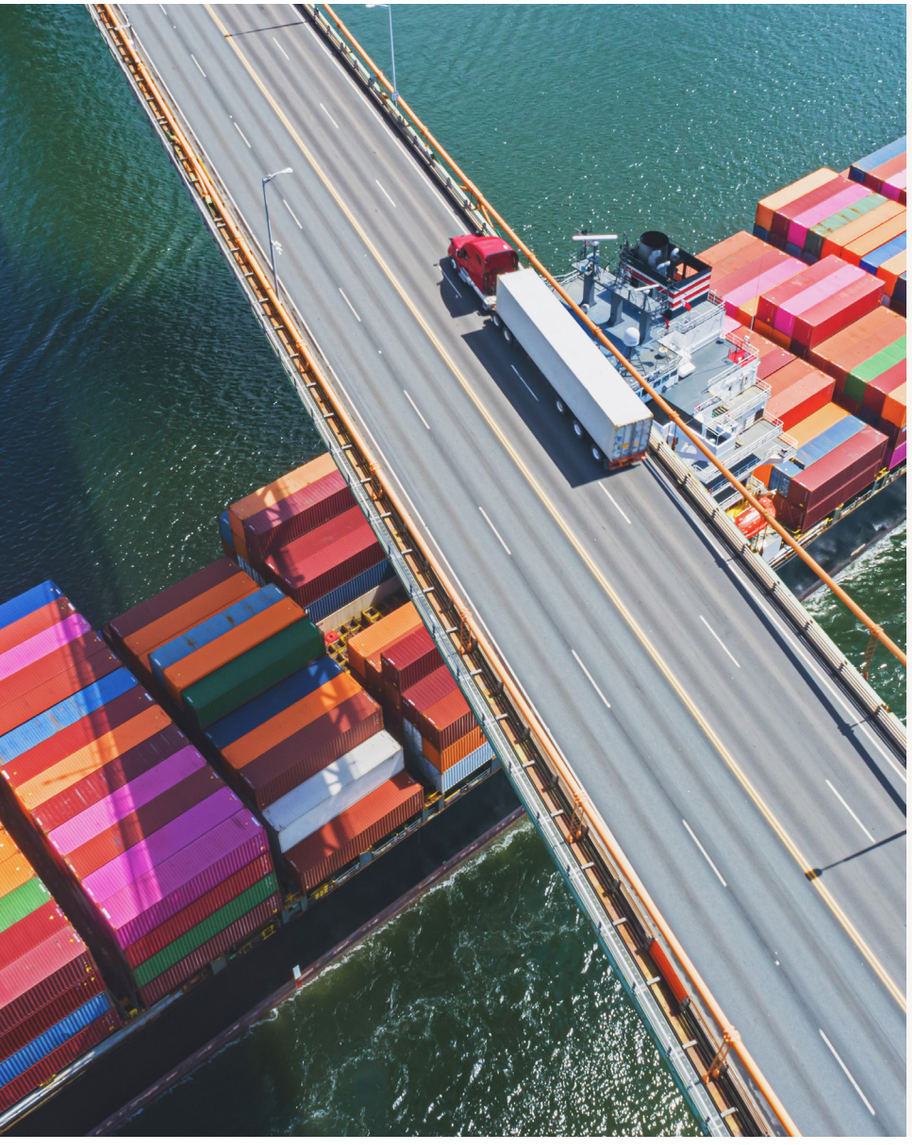


SOLUTION ACCELERATOR

Digital twins enable the safe testing of new ideas

Digital twins are virtual replications of physical manufacturing assets created using IoT-derived data. With digital representations of objects, products, equipment, people, processes, and entire manufacturing infrastructures, enterprises can perform real-time monitoring, simulation of changes, and analysis of outcomes in a controlled environment. The Databricks Solution Accelerator for Digital Twins gives organizations pre-built code, sample data and step-by-step instructions to bring digital twins to life with fault-tolerant processing of IoT streaming workloads.

- **Generate insights at scale and serve multiple downstream applications** for targeted interventions, model tuning and use case expansion
- **Test scenarios and simulate changes before implementation** to determine the most efficient course of action without wasting resources through trial and error
- **Forecast potential equipment failures and enable preventative maintenance** using real-time data to schedule services, account for outages and avoid disruptions



SOLUTION ACCELERATOR

Scalable route generation drives profitable delivery

Enterprises must deliver quickly, competitively and profitably to maintain delivery as a viable service. Scalable route generation uses IoT data to optimize navigation paths for human drivers, autonomous vehicles, drone deliveries and industrial robotics. These algorithms adapt to changing conditions like traffic congestion, detours, weather and vehicle performance to predict resource-efficient routes for individuals or fleets of devices. Enterprises can get started with the Databricks Solution Accelerator for Scalable Route Generation to do the following:

- **Easily enhance algorithms with large volumes of real-time data** from external sources
- **Quickly iterate and fine-tune models** to drive innovation with less waste, fewer resources, and at a reduced cost
- **Optimize delivery from store to customer** to narrow delivery windows and maximize margins for customer satisfaction and profitability

IIoT technologies fuel prediction with precision

The Industrial Internet of Things (IIoT) allows enterprises to design a scalable, end-to-end technology stack for the industrial application of modern IoT analytics. Most IIoT use cases aim to maximize the short-term utilization of an industrial asset while minimizing its long-term maintenance costs. Databricks Lakehouse Platform uses the power of Delta Lake to manage massive volumes of time series data alongside best-of-breed tools to jump-start enterprise use of IIoT with efficiency and speed.

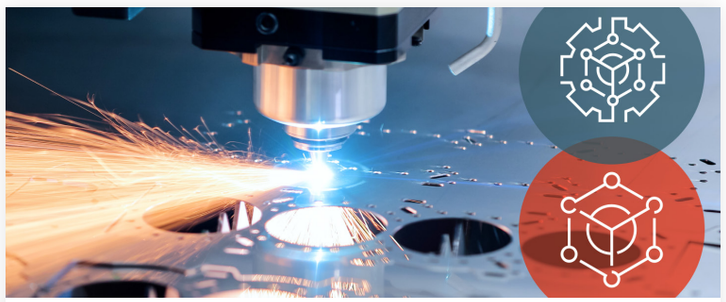
- **Ingest real-time IIoT machine-to-machine data from field devices** to build predictive maintenance models with historical data, time series shifts and custom parameters
- **Gain real-time operational analytics** on streaming time series data with complex time series processing, and use time series analysis functions for forecasting and anomaly detection
- **Fulfill machine learning use cases with IIoT data** and build downstream pipelines to enrich and aggregate IIoT application data for analytics

CHAPTER 5:

Next Steps Toward IoT-Supported Manufacturing

Learn more about how IoT data can revolutionize your manufacturing operations. Check out the following resources and meet with Databricks to explore how to best implement and leverage IoT in your environment.

OTHER RESOURCES



eBook

Making Your Digital Twin Come to Life

There's a reason why the use of digital twins among manufacturers is projected to increase by 58% per year through 2026.

Digital twins are virtual models that let you run simulations, study performance issues and generate insights you can't get anywhere else. Discover how to make them work for you in this new eBook.

Read now →

eBook

Overcoming the 3 Big Data Challenges of Industrial IoT for Predictive Maintenance

Beyond preventing equipment failures, predictive maintenance is at the core of protecting assets, improving operational efficiency and even achieving sustainability goals.

See how companies are using the Databricks Lakehouse Platform for data engineering and AI to save time, cut costs, increase safety and generate revenues from their IoT.

Read now →

eBook

Lakehouse for Manufacturing

Manufacturers generate four times more data each year than in the previous one — a faster growth rate than in any other industry. Your sensors, telematics, videos and images produce vast amounts of unstructured data full of insights for maximizing your business performance.

Learn how to harness this data in the lakehouse so you can deliver personalized outcomes, boost industrial productivity and innovate at the speed of data.

Read now →

databricks

About Databricks

Databricks is the data and AI 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.

Sign up for a free trial

databricks