Contents:

DeepIQ–Databricks Energy Solutions Partnership Overview

Executive Summary

DeepIQ is an Industrial DataOps & AI company, headquartered in Houston, that specializes in delivering intelligent data integration and contextualization for energy and asset intensive manufacturing industries. Our platform is engineered to work seamlessly with the Databricks ecosystem, enabling customers to transform complex OT and IT data into operational intelligence across upstream, midstream, and downstream environments.

At the heart of our offering is a secure, lightweight SaaS architecture optimized for Databricks Lakehouse. Through native integration with Delta Lake, Databricks APIs, and real-time ML pipelines, DeepIQ enables customers to ingest, model, and act on time series, geospatial, and relational data with domain-specific precision. Our DataStudio and CoPilot (Word Plug-In) solutions empower engineering teams to accelerate workflows in well planning, refinery optimization, and ESG reporting, all from within familiar tools and environments through our Industrial DataOps platform, powered by contextualized AI and native Databricks integration.

Key Collaboration Highlights

Unified Data Ingestion Across The Industrial Data Landscape

Leverage DeepIQ's Databricks-native ingestion pipelines to consolidate OT, IT, and enterprise data, including time series from SCADA/DCS, geospatial wellbore trajectories, and ERP records, into your Lakehouse. Built to eliminate silos and support multi-cloud flexibility, our framework ensures fast, secure, and scalable integration across all environments.

Automated Contextualization for Operational Intelligence

DeepIQ's AI-powered model builder automatically maps raw datasets into domain-specific knowledge graphs and real-time asset hierarchies, uniquely persisted in Delta Lake. We are the only contextualization platform that stores industrial graph structures natively within Databricks, enabling scalable, traceable, and AI-ready operational intelligence. This empowers engineers, not just data scientists, to build actionable digital twins, improving situational awareness in drilling, refining, and ESG operations.

DeepIQ is the only contextualization layer that leverages native Databricks capabilities to model and manage knowledge graphs and asset hierarchies directly in Delta Lake.

Real-Time Intelligence with GenAI & ML Pipelines

Deploy predictive maintenance, anomaly detection, and auto-generated technical reports using DeepIQ's CoPilot and ML Ops framework on Databricks. With seamless support for Delta Lake and Unity Catalog, our solution ensures governance, lineage, and real-time inferencing, all in your preferred workspace.

CoPilot is built on the Databricks Mosaic model, making it a natural extension of the Databricks ecosystem. This alignment enables DeepIQ to transition from Azure OpenAI to Databricks-native LLMs, providing secure, enterprise-ready generative AI. CoPilot now becomes the only Word-based documentation assistant trained directly on Databricks Lakehouse data.

DeepIQ actively uses Generative AI through CoPilot, which is powered by the Databricks Mosaic model, serving to automate documentation and surface insights from Lakehouse data

Success Stories

- 95% workflow reduction for EDW-to-cloud migration with Maverick Natural Resources
- <1 day IT-OT reporting latency in refinery operations at Global Clean Energy
- €4M+ operational gains through drilling optimization with OMV
- $10M annual impact through AI-driven field route optimization and exception handling with Maverick Natural Resources
- Enterprise-scale contextualization across 26 plants for a $15B petrochemical firm

- 85% accuracy in digitizing and mapping 1,000s of P&ID diagrams with a natural gas processor

As Databricks expands into real-time and transactional analytics, DeepIQ stands as a proven partner for co-developing and scaling industry-specific solutions. We are excited to explore joint go-to-market opportunities that leverage our complementary strengths in industrial AI and cloud-native architecture.

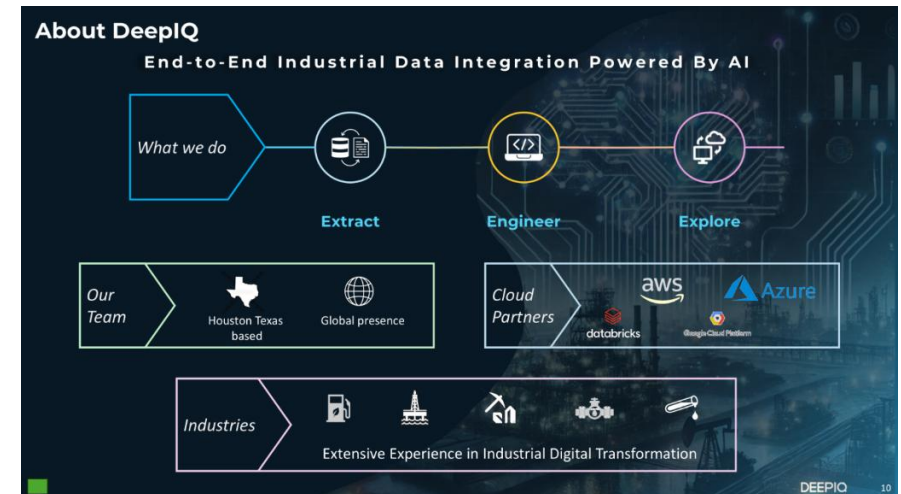## About DeepIQ – Powered by Databricks

Figure 1: DeepIQ: Industrial DataOps from Extraction to Exploration—AI-Powered, Domain-Smart, Ready to Scale

DeepIQ is an Industrial DataOps & AI company with a global presence, specializing in transforming operations for the energy and heavy asset industries through intelligent data integration and analysis.

Our platform empowers customers to extract, engineer, and explore complex data from legacy systems, real-time sources, and unstructured

files. We offer industry-leading capabilities to contextualize this data using asset hierarchies and knowledge graphs that are persisted, queried, and versioned directly in Delta Lake, an ability unique to DeepIQ in the Databricks ecosystem. These capabilities are governed, versioned, and compliant with industry standards.

By leveraging this contextualized data layer alongside Databricks Mosaic models, we automate the generation of technical documentation for compliance, design, and safety. This process ensures accuracy, repeatability, and trust in critical engineering workflows.

DeepIQ is purpose-built to work within the Databricks ecosystem. Our solutions are API-driven and integrate seamlessly with Databricks-native environments, delivering collaborative, intelligent interfaces for technical users. Whether supporting offset well analysis, field route optimization, or integrating lessons learned into design, our platform adds practical, domain-specific value to Databricks deployments in the energy sector.

DeepIQ leads the market in Industrial DataOps by unifying fragmented operational data through AI-enabled pipelines and Delta Lake-native contextualization.

Key areas of joint value with Databricks:

- Integration with Databricks Lakehouse
- Strong integration with the Databricks ecosystem, including Unity Catalog and Mosaic AI.
- Accelerators for upstream energy use cases
- Cloud-native architecture across AWS, Azure, and GCP
- Compliance-ready automation for technical documentation
- Demonstration-ready content for go-to-market campaigns

As Databricks expands its capabilities in real-time and transactional workloads, DeepIQ is well-positioned to co-develop impactful solutions that help operators act on data with clarity, speed, and confidence.

DeepIQ's Lightweight SaaS Architecture Integrated with Databricks

This architecture is the backbone of DeepIQ's Industrial DataOps capabilities—delivering governed, scalable AI pipelines for real-time operational insights.
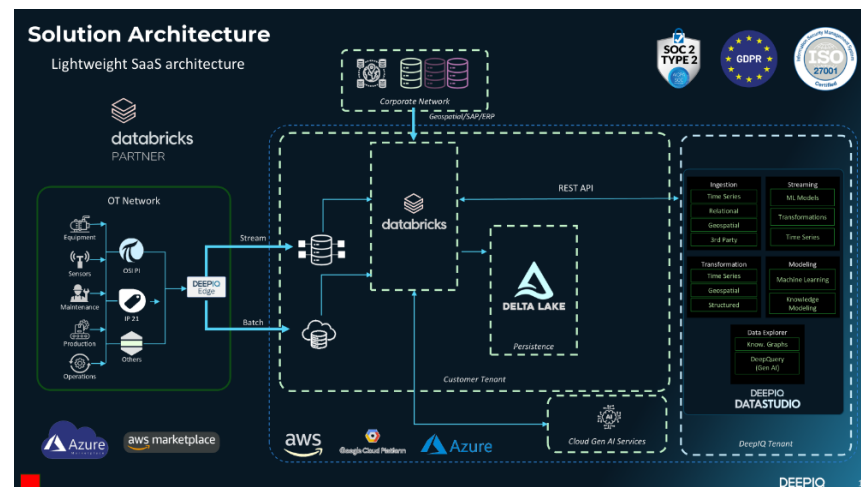

Figure 2: Industrial DataOps at Scale: Secure, Streaming, and AI-Ready from Edge to Insight

DeepIQ delivers a secure, cloud-native architecture purpose-built for industrial data environments. The solution is designed to operate across complex OT and IT networks, enabling real-time and batch data processing from field equipment to advanced analytics at scale. At its core, this architecture is optimized for Databricks Lakehouse, leveraging Delta Lake for persistent, high-performance data management.

The system starts at the OT network edge, where DeepIQ connects to a wide range of data sources, including OSI PI, IP21, and other industrial protocols. Through the DeepIQ Edge component, data is collected and streamed or batched into the processing environment hosted on

Databricks. This architecture supports integration with both cloud and corporate systems, including ERP, geospatial, and SAP platforms.

Once inside the Databricks environment, data is persisted in Delta Lake, where it becomes available to both DeepIQ's pipeline services and Databricks-native tools. Data is processed through a series of ingestion, transformation, and modeling stages supported by DeepIQ's DataStudio. This layer enables use cases such as:

- Ingestion from time series, relational, geospatial, and third-party sources
- Streaming pipelines for real-time ML inference and transformation
- Transformation and enrichment of structured and unstructured data
- Advanced modeling using machine learning and domain-specific knowledge graphs
- Natural language interfaces powered by DeepQuery (Gen AI) for technical search and exploration

All services operate within secure cloud environments, deployable in AWS, Azure, or Google Cloud, and support complete tenant separation between customer environments and DeepIQ services. The architecture complies with SOC 2 Type II, ISO 27001, and GDPR standards, ensuring enterprise-grade security and robust data governance.

Together with Databricks, this architecture empowers industrial customers to rapidly build and scale intelligent applications that convert raw data into operational insights with minimal overhead and maximum flexibility.

To strengthen our collaboration, we are enabling CoPilot to support Mosaic Model Serving. This allows us to embed CoPilot directly within Databricks-native applications, such as Genie, enabling secure, governed, and scalable GenAI across documentation workflows.

DeepIQ Word Plug-In (CoPilot): Industrial DataOps Assistant for Technical Documentation
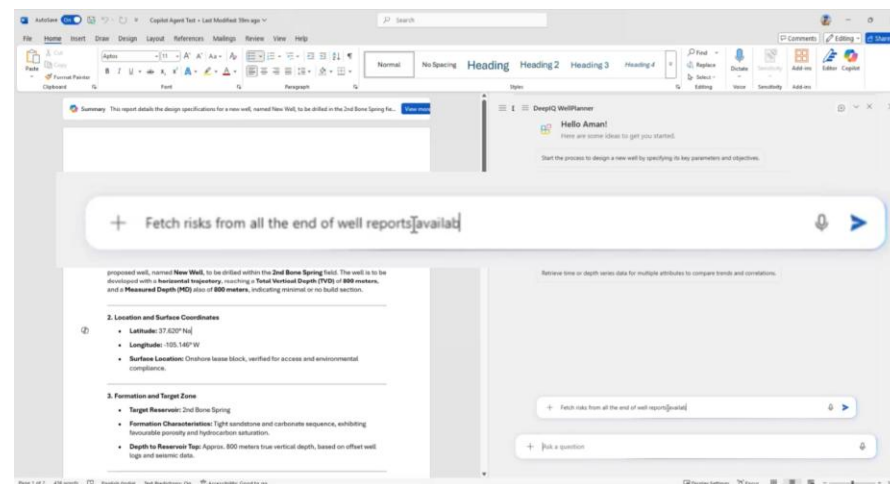

Figure 3: Ask. Analyze. Automate. CoPilot Brings Industrial DataOps into Word

The DeepIQ Word Plug-In, known as CoPilot, is a domain-specific AI assistant purpose-built for drilling and well engineering teams. Integrated directly into Microsoft Word, CoPilot connects users to DeepIQ's industrial knowledge graph and Databricks-native data pipelines, automating the generation of high-quality technical documentation, including drilling data, offset analyses, risk assessments, and post-well reviews.

Unlike general-purpose AI tools, CoPilot is powered by curated, structured data. It leverages offset well analytics, structured time series, historical drilling reports, and best practices across regions to produce documents that are technically sound, context-aware, and compliant with engineering standards.

What CoPilot does for drilling teams:

- Surfaces zone-specific offset insights like stuck pipe events, loss circulation zones, and mud weight windows
- Queries DeepIQ's knowledge graph for similar wells to extract lessons learned, typical hazards, and recommended mitigations
- Auto-populates document sections based on formation targets, casing designs, and operational constraints
- Enables engineers to ask natural-language questions inside Word and receive validated responses from real operational data
- By embedding AI into documentation workflows, CoPilot extends Industrial DataOps to the desktop, linking engineers directly to governed operational data.

How CoPilot works with Databricks:

Enhanced Description of DeepIQ's Co-pilot Integration with Databricks

DeepIQ has pioneered the persistence of industrial knowledge graphs and asset hierarchies within Delta Lake, not just for drilling, but across domains, providing a foundation for contextualized reasoning, technical automation, and enterprise-scale digital twins natively in the Databricks environment. Leveraging specialized R&D to achieve a clear separation between schema management and the persistence layer. This architecture enables scalable, flexible knowledge representation for complex industrial data.

How Co-pilot Leverages Databricks Mosaic Model Serving

DeepIQ's Co-pilot utilizes Databricks Mosaic Model Serving to deliver advanced automation and intelligence across several critical functions:

- **Document Parsing**: Automatically extracts and interprets content from technical documents.
- **Context Generation**: Creates rich, contextual understanding of technical content for more accurate analysis.

- **Knowledge Graph Retrieval**: Seamlessly accesses and utilizes data from knowledge graphs stored in Delta Lake.
- **Content Reformatting**: Adapts and restructures information to meet specific documentation standards and requirements.

Seamless User Experience

Co-pilot is fully compatible with Databricks Genie and similar front-end solutions, providing technical teams with a unified, end-to-end workflow, from data exploration to automated documentation. This integration ensures a smooth, intuitive experience for users working with complex engineering data.

Extensibility Beyond Drilling

While Co-pilot is currently deployed within the drilling domain, its architecture is designed for adaptability. The solution can be extended to support any technical documentation process that benefits from integrating insights across disparate data sources, making it a versatile tool for a wide range of industrial and engineering applications.

CoPilot on Databricks Genie

CoPilot can be extended into Databricks Genie to deliver an immersive, collaborative experience for drilling and engineering teams directly within the Databricks environment. By embedding CoPilot's natural language interface within Genie, users can interact with DeepIQ's knowledge graph and offset analytics using chat-style queries, retrieve structured content, and generate engineering documentation all within the Databricks workspace.

This integration would enable engineers to explore historical well data, surface relevant lessons learned, and pre-fill drilling program templates in real-time, all within a secure, Databricks-native interface. Genie would act as the front-end layer, while DeepIQ's APIs and Delta Lake-backed knowledge models power the data retrieval, contextualization, and

document generation behind the scenes. This creates a seamless, domain-aware workspace for energy professionals to analyze complex data and accelerate high-quality decision-making.
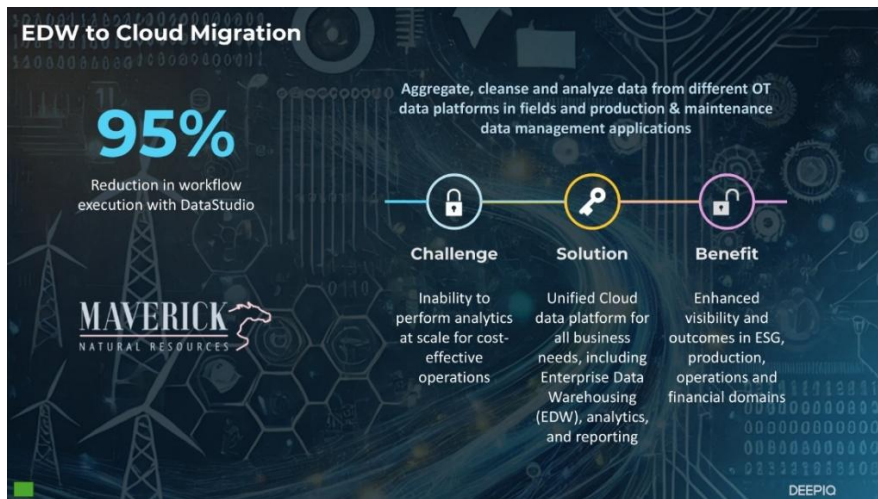
Additional Use Cases on Databricks

Figure 4: Field DataOps for Real-Time Route Optimization

DeepIQ's EDW to Cloud Migration solution, as deployed with Maverick Natural Resources, operates seamlessly with Databricks as the central processing and analytics engine. By using Databricks as the unified platform for data ingestion, transformation, and warehousing, DeepIQ enables scalable and high-performance workflows that consolidate OT and IT data from production, maintenance, and ESG systems.

DataStudio orchestrates these pipelines across Databricks' Lakehouse architecture, allowing real-time analytics and automated reporting for cross-functional business domains. This architecture has delivered a 95%

reduction in workflow execution time while dramatically improving visibility across operations, finance, and sustainability initiatives.
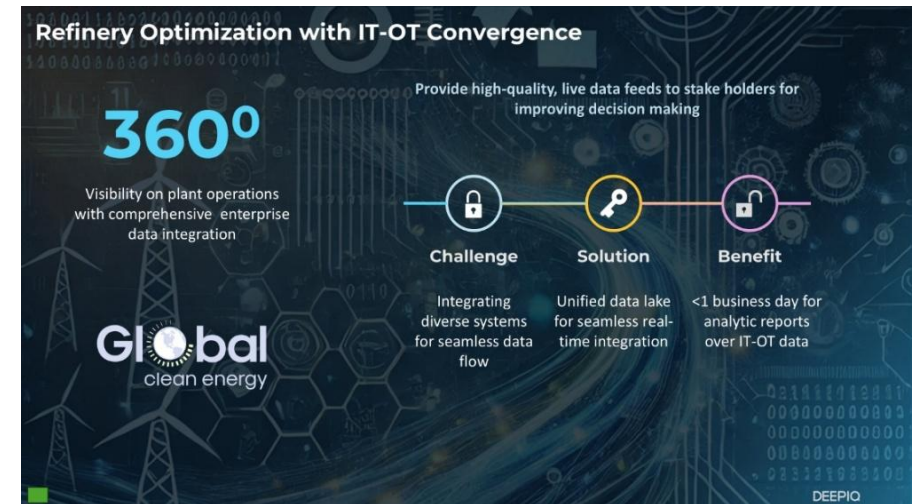

Figure 5: Industrial DataOps for OT/IT Convergence in Refining

DeepIQ's refinery optimization solution for Global Clean Energy relies on Databricks to unify IT and OT data sources into a single, real-time analytics environment. Using Databricks' scalable data lakehouse architecture, DeepIQ integrates diverse systems across the plant, from control systems and sensors to enterprise IT platforms, enabling continuous, high-quality data flows into a centralized environment.

This seamless integration supports near real-time reporting and analysis across operations, maintenance, and business domains. By orchestrating ingestion and transformation pipelines within Databricks and delivering insights through DeepIQ DataStudio, the solution reduced reporting latency to under one business day, providing 360-degree visibility for faster, data-driven decision-making across the refinery.
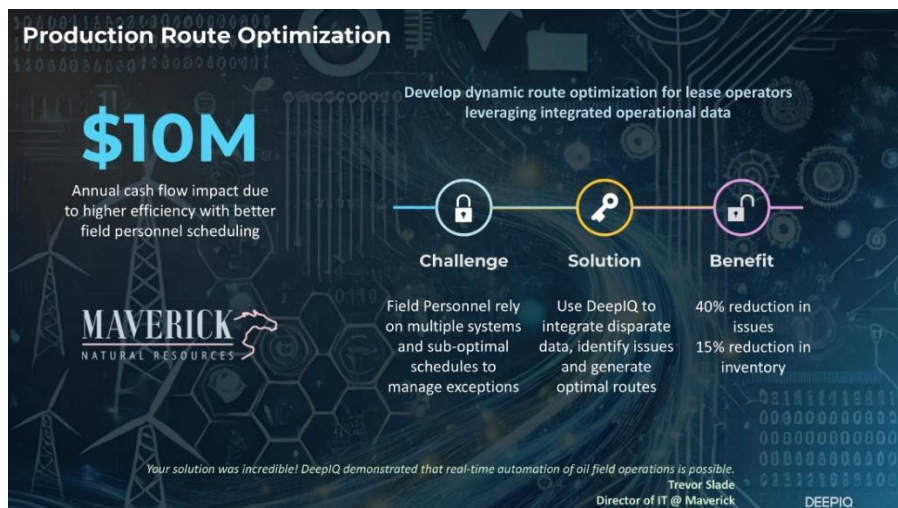
Figure 6: AI-Powered Route Optimization Built on Industrial DataOps

Figure 7: DataOps-Powered Simulation for Drilling Optimization

Maverick Natural Resources partnered with DeepIQ to optimize lease operator routes using real-time operational data. By integrating disparate field systems and time series data into the Databricks Lakehouse, DeepIQ enabled dynamic scheduling and automated exception handling. Field personnel, who previously relied on static, suboptimal schedules across disconnected systems, gained visibility into current field conditions and could respond proactively. Using AI-driven route generation and exception management workflows,

Maverick achieved a $10M annualized cash flow impact, a 40% reduction in operational issues, and a 15% reduction in field inventory. This project demonstrated how DeepIQ's platform, operating natively on Databricks, can automate oilfield operations at scale while delivering measurable operational gains.
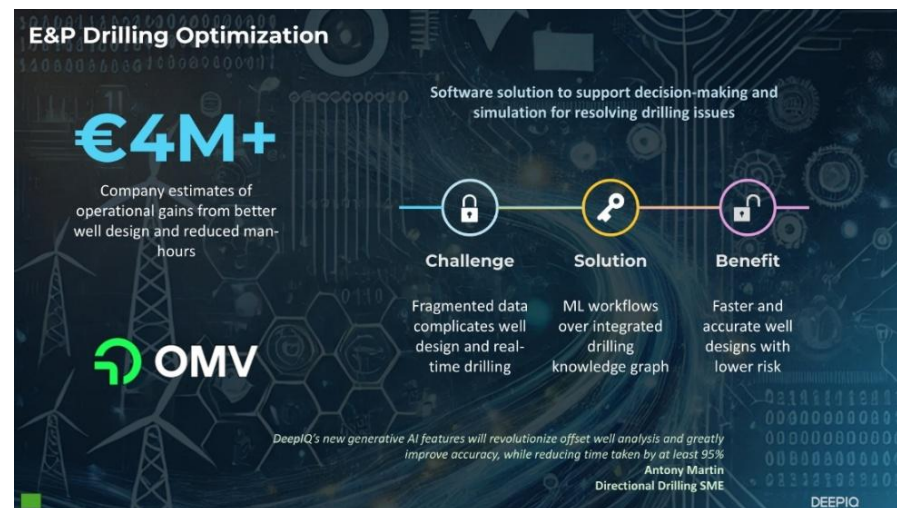
DeepIQ partnered with OMV to optimize drilling operations using a Databricks-powered architecture that integrates fragmented operational data into a unified, ML-driven environment. At the core of the solution is a drilling knowledge graph built on data ingested and transformed through Databricks Delta Lake, enabling scalable analytics and machine learning workflows.

By combining offset well data, real-time drilling parameters, and historical lessons learned, the solution supports advanced simulation and decision-making during well planning and execution. This integration led to over €4 million in estimated operational gains, driven by faster and more accurate well designs, reduced risk, and significantly lower engineering hours. Databricks served as the high-performance engine for data processing, modeling, and knowledge retrieval across the entire workflow.
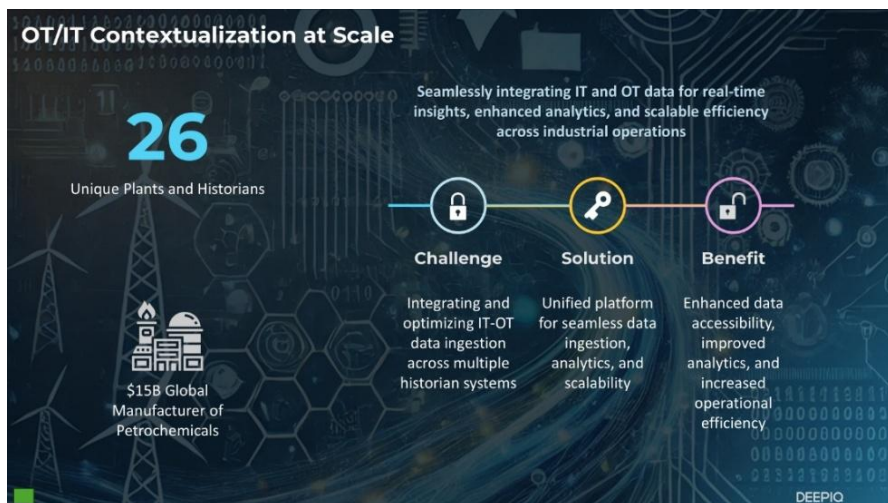
Figure 8: Scaling Industrial DataOps Across 26 Plants

Figure 9: AI-Driven Contextualization in Industrial DataOps Workflows

DeepIQ is currently partnering with a $15 billion global petrochemical manufacturer to scale OT/IT contextualization across 26 unique plants and historian systems, utilizing Databricks as the central data and analytics engine. The solution consolidates diverse sources of time series and enterprise data into a unified platform that enables seamless ingestion, transformation, and analytics.

Built on the Databricks Lakehouse architecture, this deployment supports both real-time and batch processing, enabling the unification of operational and business insights at scale. By leveraging Databricks for scalable compute and DeepIQ DataStudio for domain-specific analytics and modeling, the customer has significantly enhanced data accessibility and operational visibility. The project is already delivering measurable improvements in data quality, reporting efficiency, and system-wide analytics across global assets.
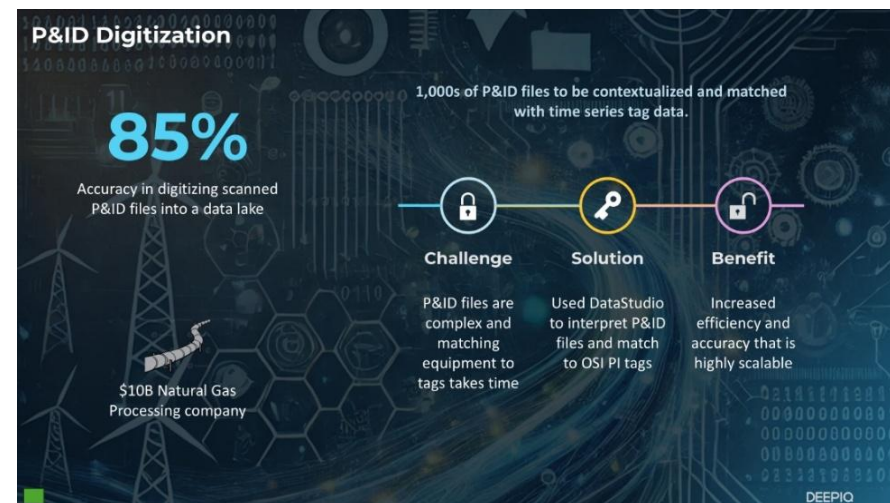
DeepIQ collaborated with Databricks on a joint initiative to digitize scanned P&ID diagrams for a $10B natural gas processing company. The project aimed to extract and contextualize equipment data from complex engineering drawings and align it with real-time OSI PI tag data. Using DeepIQ DataStudio and Databricks' scalable Lakehouse platform, the team created a pipeline that interprets image-based P&ID files, identifies key entities, and links them to operational time series data.

This approach achieved 85% accuracy in file-to-tag matching while dramatically improving the speed and scalability of asset contextualization. Databricks served as the foundation for data persistence, processing, and model execution, enabling the operationalization of insights across plants with consistent accuracy and minimal manual intervention.