Cloudwick

Amorphic

Transform Your Industry 4.0 Journey with a Modernized Manufacturing Data Supply Chain

Build Data Unification and Machine Learning Capabilities Using the Amorphic Data Cloud for AWS

Summary Abstract

Manufacturing organizations benefit from the automation brought by the Third Industrial Revolution, but struggle to extract value from and manage the massive volumes of data created by 20th Century process innovation. The Fourth Industrial Revolution (Industry 4.0) - now present - promises to offer tools to unlock the value of these massive quantities of data with the use of technologies such as Data Analytics, Machine Learning & Industrial IoT. However, manufacturing companies still struggle to convert this data value into business outcomes due to the proliferation of data silos, and the overwhelming prevalence of complex and unwieldy data supply chains which leaves data in an unmanageable state.

This paper explores:

a. How manufacturing companies benefit by modernizing their data supply chain with a data cloud.

b. How a modern data supply chain can help improve Overall Equipment Effectiveness (OEE).

c. How an Industry 4.0 transformation use case with a purpose-built data cloud results in increased innovation and improved OEE.

About Amorphic

Cloudwick's Amorphic Data Cloud for AWS helps organizations gain deeper insights from their data. Amorphic makes any data sharable, searchable, and analyzable by any analytic user, data scientist or any data-driven application developer. Encompassing a built-in data catalog for search and sharing, enterprise-grade governance and security, along with machine learning (ML) and artificial intelligence (AI) capabilities for advanced analytics, Amorphic provides a production ready, self-service data cloud for all data and all workloads, natively integrating over 65 services on AWS with "single pane of glass" simplicity for better data-driven decision making with near zero IT development or support required.

About Cloudwick

Cloudwick powers more agile, innovative, and cost-effective cloud data and analytics solutions. Cloudwick's Amorphic Data Cloud for AWS makes all data searchable, shareable, and analyzable by any user for any workload, enabling organizations to gain deeper insights from their data. Cloudwick's Professional Services provides organizations with trusted, high-impact data and analytic migration and modernization outcomes. Founded in 2010 in Newark, CA, Cloudwick has over 210 technology professionals with 400+ Certifications serving customers throughout the US, EMEA, and APAC. Cloudwick offerings are available direct, through the AWS Marketplace and through select resellers.

A Brief History of Industrial Automation in Manufacturing

The story of industrial automation and process standardization dates to the First Industrial Revolution in the eighteenth century, when steam engines mechanized weaving looms to speed production. A century later, the Second Industrial Revolution introduced mass production lines powered by electricity to speed production of everything from pins to widgets to clocks. The Second Industrial Revolution also witnessed the expansion of electricity, petroleum, and steel. In 1913, Ford Motor Company introduced a car production assembly line, which pioneered automation in the manufacturing industry. The Third Industrial Revolution, or Digital Revolution, began in the late 1900s, with the development of electronics and computers to automate work and to increase the speed of data processing.

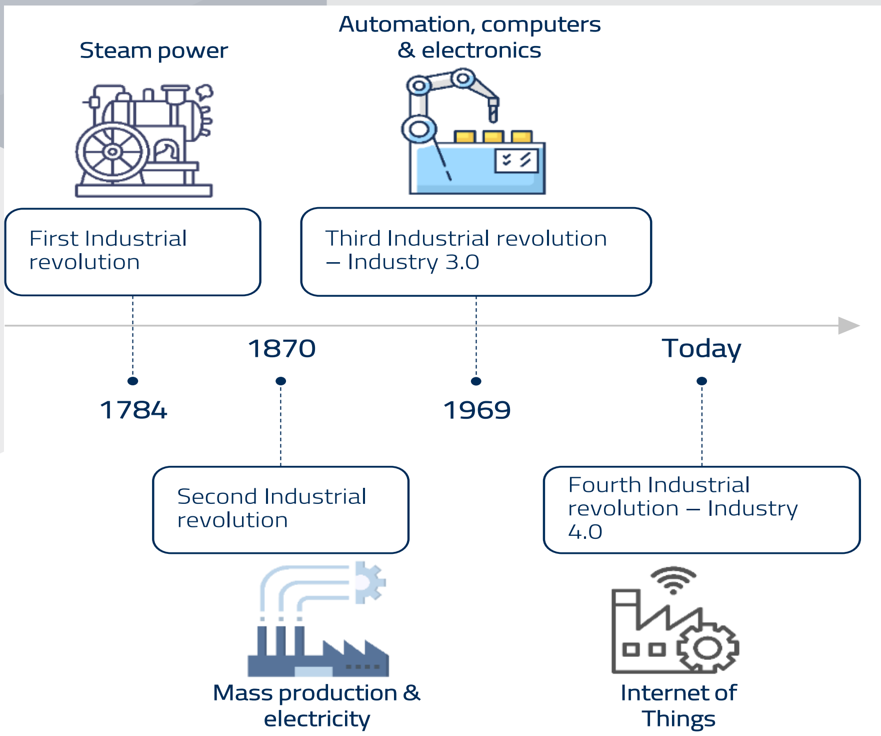

Exhibit 1 - Timeline of Industrial Revolutions

Most industrial operations automate to boost production output and quality, and to reduce the cost of labor. A manufacturing organization that uses the latest technologies to automate its processes typically sees improved throughput and efficiency, higher-quality products, reduced labor requirements, and lower production costs.

Digital transformation, data analytics, artificial intelligence (AI), and Industrial IoT (IIoT) mark the current era of the Fourth Industrial Revolution, often called Industry 4.0. In the next section, we explore the value promise of the Industry 4.0 transformation to provide the benefits of industrial automation for the Digital Age.

The $3.7T Value Promise of Industry 4.0 Technologies—The Benefit of Modernization and Potential Roadblocks

As per a study performed by McKinsey and the World Economic Forum[1], the Fourth Industrial Revolution (Industry 4.0) is likely to create up to $3.7 trillion in value by 2025. Compared to the Third Industrial Revolution, the Fourth Industrial Revolution (Industry 4.0) will have a relatively high impact on production output with comparatively little replacement of equipment. Unlike prior industrial revolutions, Industry 4.0 is not about replacing the existing technology assets with new ones but about mastering the disruptive technologies along three different dimensions[1]: operational effectiveness, digital transformation, and new business models.

During the Industry 3.0 automation period, Information Technology (IT) & Operational Technology (OT) stacks were built independently of each other. This system isolation, or data siloing, has been the source of a great deal of manufacturing inefficiency that still needs to be overcome. Industry 4.0 aims to remove the operational inefficiencies caused by system isolation through converging the IT and OT stacks. Much of the gain from reducing data silos is expected in the areas of productivity, cost reduction, efficiency, and quality.

Listed below are eight critical value drivers that will help realize these gains across the manufacturing value chain. Improving these value drivers will be critical to achieving Industry 4.0 transformation.

1 Resource/Process	2 Asset Utilization	3 Labor	4 Inventory
Optimize material consumption, improve speed/yield	Maximize for best use of company's machinery	Optimize this important cost driver for companies	Reduce capital tied up to inventory
5 Quality	**6 Supply/Demand**	**7 Time-to-market**	**8 Service**
Maximize quality to reduce the extra cost of scrap and rework	Maximize value captured from the market by matching supply and demand	Gain early mover advantage	Reduce cost of operations

Exhibit 2 below further quantifies these eight value drivers powered by IT & OT convergence as per the McKinsey study previously cited.[1]

[1] The Next Economic Growth Engine Scaling Fourth Industrial Revolution Technologies in Production January 2018 The World Economic Forum in collaboration with McKinsey & Company

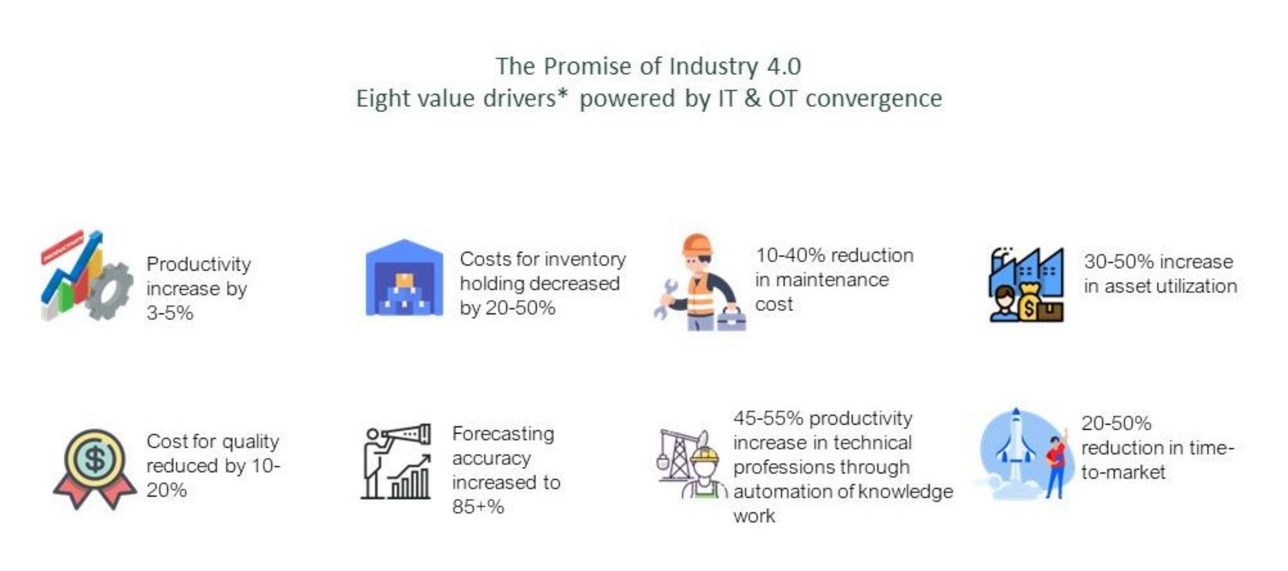

Exhibit 2: The Value Promise of Industry 4.0[2]

IIoT & Data are the key technology enablers for the IT/OT convergence required to fuel the above value drivers. IIoT helps convert the physical attributes of every manufacturing element into digital format using sensors. Data Analytics and Machine Learning helps process and analyze this data and turn it into business outcomes. As plant floors are digitized with IIoT sensors, more data gets generated, making data management and analysis a holy grail of Industry 4.0 transformation: that is, converting raw data into meaningful information which can improve business outcomes. However, managing these enormous volumes of data throughout what is termed the "data supply chain" can be a difficult and unwieldy task.

There are high expectations from Industry 4.0 technologies to drive the Fourth Industrial Revolution. Yet only about 30 percent of companies are capturing value from Industry 4.0 solutions at scale today. In the remainder of this paper, we examine how the data supply chain can be major roadblock for Industry 4.0 adoption and explore how to address and overcome these roadblocks.

[2] "Industry 4.0 How to navigate digitization of the manufacturing sector" McKinsey Digital 2015 page 2

Complex and Siloed Data Supply Chains Limit Your Ability to Pursue Industry 4.0 Transformation

Data hides in every nook and cranny of Enterprise legacy applications and databases. Siloed data necessarily impedes an organization's ability to derive cohesive insights from data and can limit the benefit of Industry 4.0 transformation. To alleviate this issue, organizations today seek a well-oiled *data supply chain*: an automated and securely governed process that can 1) ingest and consolidate data of all types – structured, semi-structured, unstructured - from all sources into a single location to be stored and processed, 2) transform the data to allow analysis and extraction of data insights, and finally 3) a process that facilitates data visualization and data sharing throughout an organization so stakeholders can make better business decisions.

Let's look at a sample typical data supply chain in a manufacturing organization:

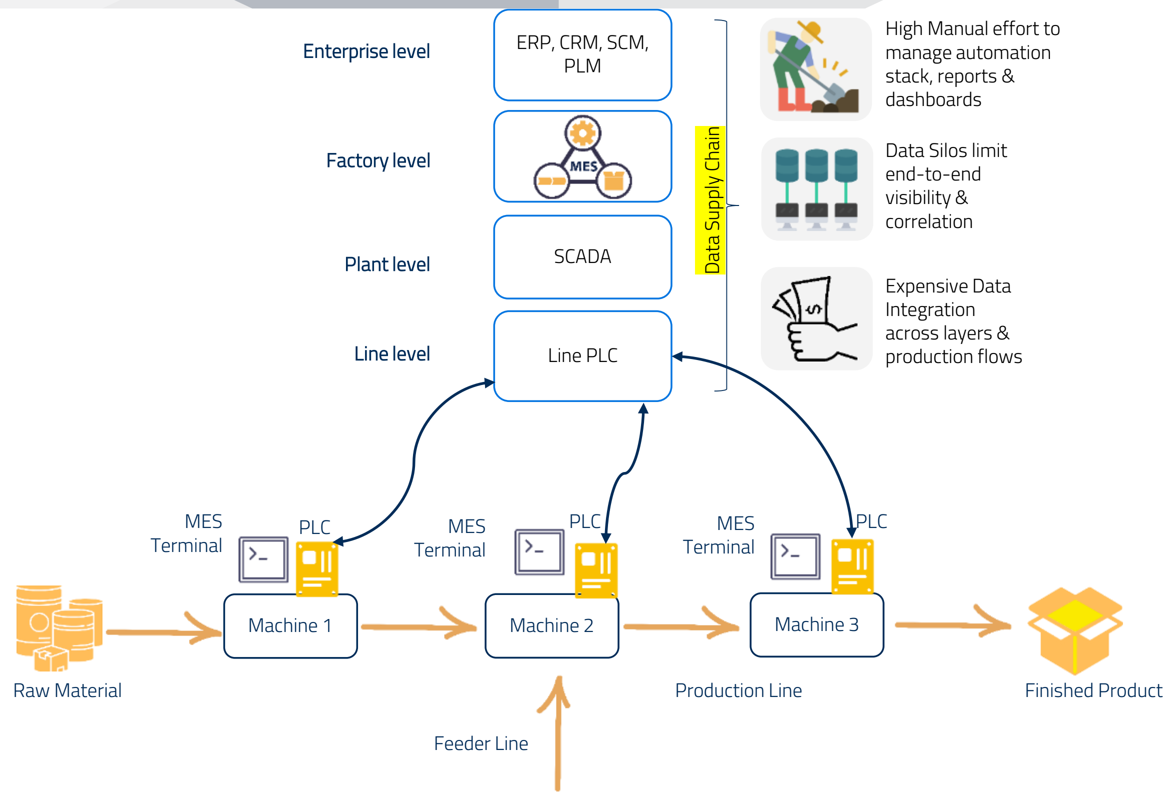

Exhibit 3 - Five Layered Industrial 3.0 Automation Stack

Exhibit 3 shows a typical 5-layered industrial automation stack having production lines with machines supplemented by feeder lines. Each machine has a PLC attached which feeds real-time health of the machine (machine telemetry data) to the line level PLC, which rolls this data into the plant level SCADA system. This system facilitates supervisory access and control of all the machines and production lines in the plant.

The top-layer of this 5-layered industrial automation stack are the Enterprise-Level functions such as ERP, CRM, SCM, and PLM. The MES layer automates execution of work orders at the factory level spanning across multiple machines in the production line. Another important function of the MES layer is to calculate OEE (Overall Equipment Effectiveness) which is considered the gold standard in measuring availability, productivity, and quality.

When built, each layer was originally designed to serve a specific purpose, and later cobbled together with other layers using expensive and time-consuming mechanisms. In most cases, there is no architectural integration between layers in the automation stack.

This lack of architectural integration creates common hurdles faced by any manufacturing company trying to implement a smooth data supply chain in their Industry 4.0 pursuit. These hurdles include:

1. Data Silos - data trapped behind application silos that limit the ability to extract data for analysis
2. Disparate Automation Layers - requiring expensive integration with third- party equipment and human effort that limits the scaling of data supply chains
3. Manual Paper-Based Production Flows – data and information it typically not digitized
4. Limited Visibility into Operations - lack of process transparency due to items #1-3

The above problems – illustrated on the right-side of the Exhibit 3 – create a lack of data sharing between layers of the automation stack, resulting in very complex and fractured data supply chains. A fractured, complex data supply chain restricts the ability to democratize data access across the organization and forestalls data insight sharing across manufacturing operations.

The creation of a functional, fluid data supply chain doesn't have to be an enigma. In the next section, we investigate how manufacturing companies can modernize their data supply chains as a first step in their Industry 4.0 journey.

Modernize Your Manufacturing Data Supply Chain with a Data Cloud Providing Data Unification & Machine Learning

An HBR article "Your Data Supply Chains are Probably a Mess - Here's How to Fix Them"[3] highlights how companies that employ 'data supply chain management' report better results in obtaining data insights. A modern streamlined data supply chain handles all phases of data management, i.e., collection, organization, and consumption of data. This type of data supply chain can be established using the same principles of process and quality management used in physical supply chains.

Exhibit 4 helps explore this potential; it shows a three-part data supply chain management using a Data Cloud*:

1. Automatic ingestion from individual layers of the automation stack in batch and real-time.
2. Cloud data unification/centralization & access to advanced analytics/machine learning as provided by a Data Cloud.
3. Consumption of data to derive insights and create products fueled by data and machine learning to solve problems.

**A Data Cloud is a self-service, production ready SaaS offering that enables manufacturing organizations to unify and connect all their siloed data into a single source of truth for analysis and data insight extraction. A Data Cloud typically provides secure, governed access to advanced analytics/machine learning capabilities which allow users – from analytic users to data scientists to data-driven developers or line of business leaders - to analyze, and share transformed data and the resulting data insights across an entire organization.*

[3] "Your Data Supply Chains Are Probably a Mess. Here's How to Fix Them" by Tom Davenport, Theodoros Evgeniou, and Thomas C. Redman, HBR article June 24, 2021

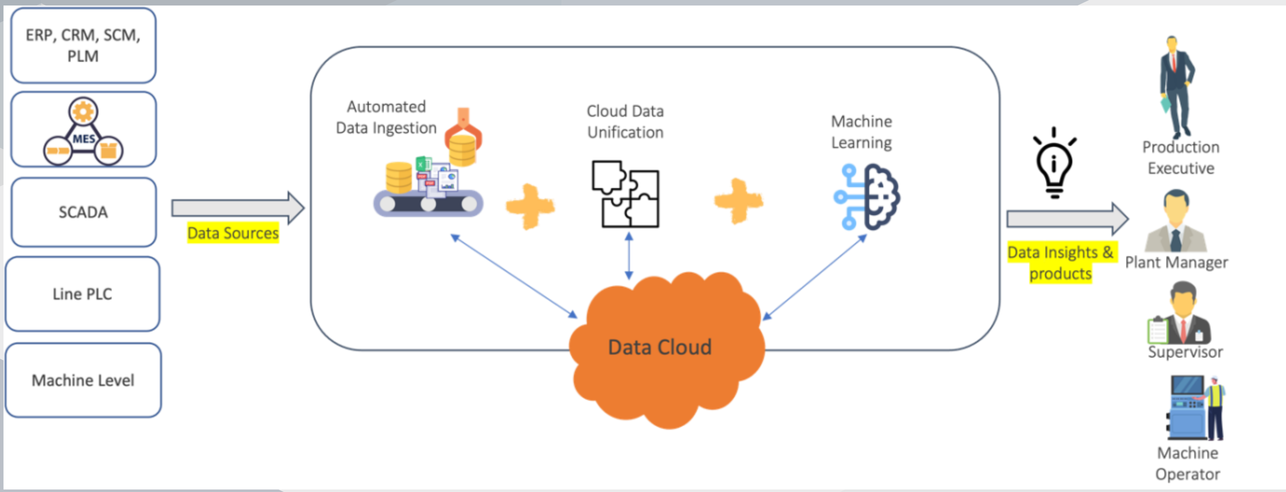

Exhibit 4: A Manufacturing Data Supply Chain Using a Data Cloud

In this scenario, utilizing a data cloud to consolidate formerly siloed data into a single location and then to apply machine learning to data for extract higher quality insights, leads to beneficial outcomes: a machine operator using a machine learning data product to maximize production run performance or a maintenance supervisor consulting predictive maintenance data to anticipate and alleviate an otherwise probable machine failure.

In the next sections, we explore into how a streamlined data supply chain using a data cloud can help improve metrics like OEE.

How Can a Modern Data Supply Chain Improve OEE?

OEE (Overall Equipment Effectiveness) is the gold standard for measuring manufacturing productivity, as it identifies the percentage of manufacturing time that is truly productive. Measuring OEE to identify underlying availability, performance, and quality losses can provide meaningful insights into the causalities contributing to each loss. Manufacturing companies can use these insights to drive quality improvement programs.

Exhibit 5 shows the relationship between OEE calculation, different types of losses, and the corresponding causal factors shown on the right.

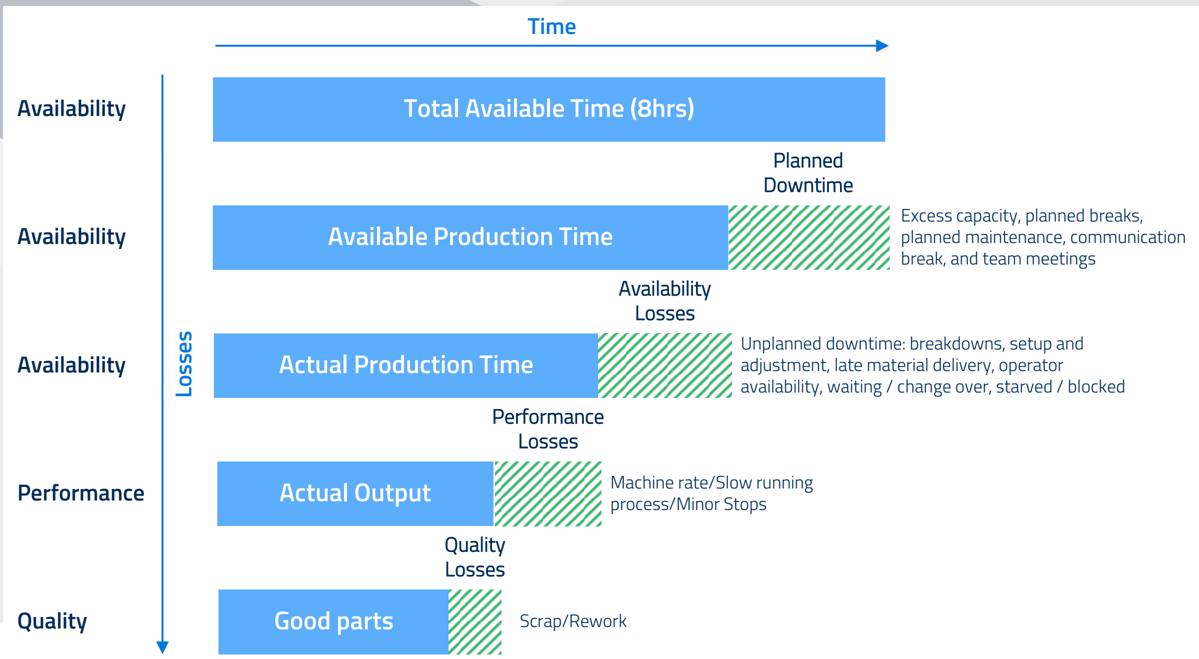

Exhibit 5: OEE calculation with production losses and its causalities

Data hidden in different layers of the automation stack holds key insights capable of detailing the reasons for losses. A modern data supply chain brings all data from each of the stack layers into a centralized data cloud, as shown in Exhibit 4. Data unification and correlation of data across layers will provide insights that help in improving OEE. Automating this activity identifies issues and creates a resolving mechanism to improve OEE. These insights help make continuous improvements possible throughout the manufacturing process, which increase OEE.

Further, a modern data supply chain leverages machine learning to learn from data sources, and then to begin predicting OEE and providing recommendations on how to improve it further. Machine learning can process enormous volumes of data and perform correlations between casualties and OEE at an unprecedented scale beyond human ability. Exhibit 6 stepwise illustrates how data unification & machine learning in a data cloud help improve OEE.

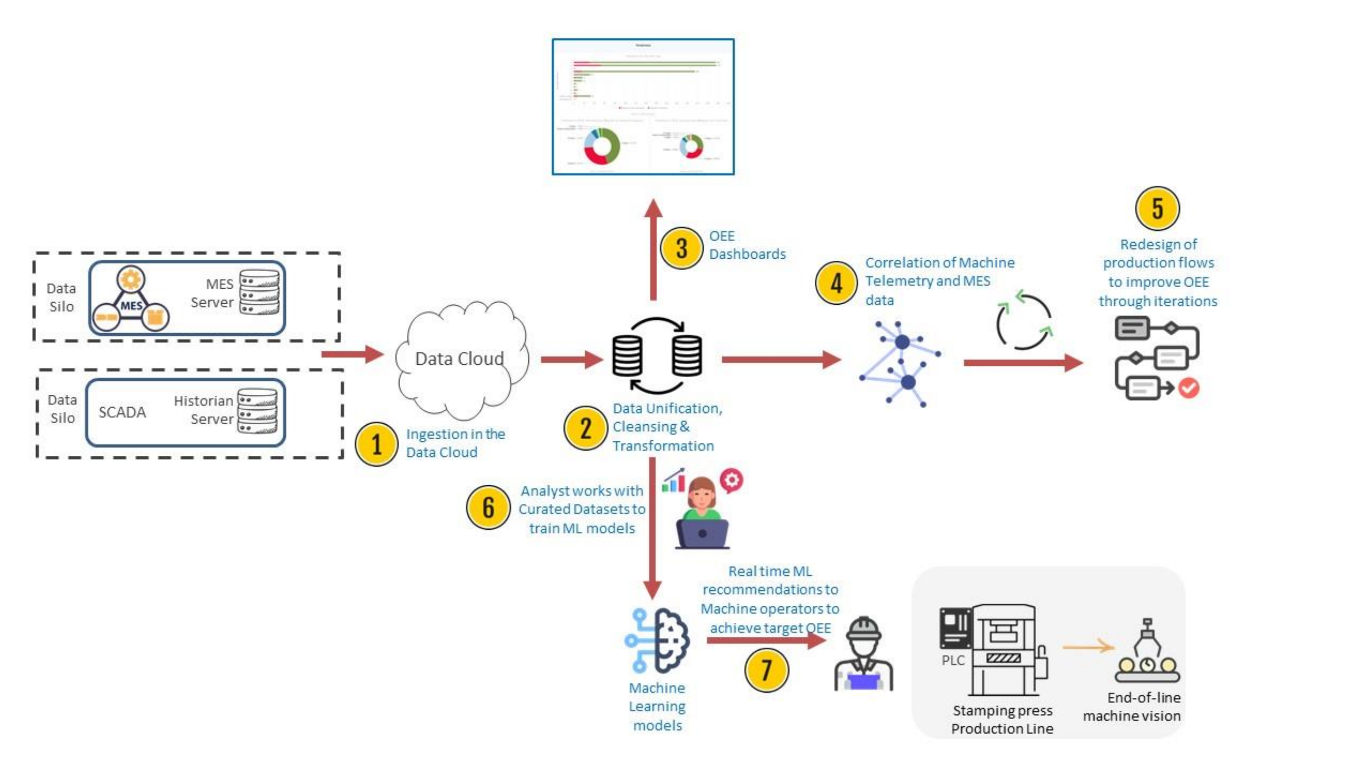

Exhibit 6: How Data & Machine Learning Improve OEE

In the next section, this paper explores an applied use case of data unification and machine learning in an industrial setting.

Industry Use Case - Preventing Quality Defects & Predicting Through-Put with Data Unification & Machine Learning in a Data Cloud

Business Problem:

A stamping press production line used for manufacturing aluminum cans from a sheet of aluminum faces a problem with a high yield loss of 10%, with an OEE of 60%. Management believes that a 1% improvement in OEE will yield an additional capacity of 1 million cans. Of the 10% quality yield loss, there is limited visibility into the defects causing the loss. Management suspects that a proper loss detection mechanism could recover good cans and improve OEE, but data is in silos and difficult to analyze without onsite experts.

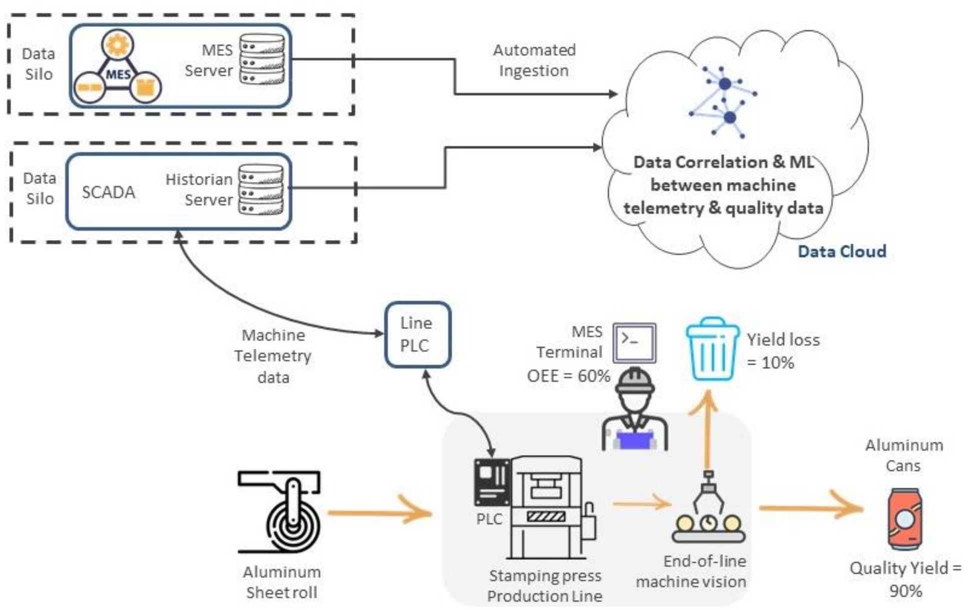

Exhibit 7 - Improving Quality Losses with Data Unification and Machine Learning in a Data Cloud

Technical Challenges:

1. How to break down silos and transition data and analytics to a centralized data cloud.
2. Lack of data engineering skills and modern data analytic expertise.
3. Lack of visibility - require automation of OEE dashboard observability.
4. SCADA systems not collecting clean data.
5. Current OEE systems and their data are siloed and not available for correlated analytics.

Desired Goals:

1. Expand existing OEE improvement capabilities with a centralized data repository such as a cloud platform.
2. Improve human decisions with machine data and share data transparently across teams.
3. Increase Aluminum Can yield to drive higher profitability from the current manufacturing process. Every 1% OEE improvement yields an additional capacity of 1 million cans.

Solution Use Cases:

Use Case	Description
BI	OEE dashboards with Stamping Press & Quality Yield by SKU
Analytics	Correlation of Machine Telemetry and Quality Data
ML	Reduce False Failures by Offline Labeling of Defects Fed Into ML
Prescriptive Analytics	Preventing Quality Defects and Predicting Throughput

Working with a Turn-Key Solution: Cloudwick's Amorphic Data Cloud

In this case study, the organization chose to implement Cloudwick's Amorphic Data Cloud for AWS, a data cloud solution that natively integrates over 65 AWS Data & Analytics and other services into a single simplified user interface. The Amorphic Data Cloud offered this customer a production-ready, self-service AWS ISV Certified solution to help them quickly ingest, transform, secure and govern, find, share, use and visualize data using the AWS Modern Data Analytic stack.

Amorphic enabled them to migrate data and modernize data analytics without having to build or support the solution. Amorphic offered a turn-key solution that was up and running within 90 minutes directly within the customer's AWS Account.

The steps to implement the solution were straightforward:

Implementation tasks:

1. Set-up Amorphic Data Cloud for OEE in customer's account.
2. Identified OEE data sets and developed KPIs (availability, performance, quality, OEE, machine telemetry correlation and vision system rejects) with stakeholders
3. Ingested MES, Historian & Vision System Data Sets to Amorphic Data Cloud.
4. Performed quality and validation checks on all data sources. Cloudwick team highlighted observed data gaps to the relevant stakeholders.

5. Built data transformation jobs to automate correlation from different systems with OEE.
6. Developed OEE dashboard for KPI and correlation visualizations.
7. Performed a series of Machine Learning experiments on the OEE & Machine Telemetry data.

The above tasks provide implementation level details of the steps given in Exhibit 6 for improving OEE using data & machine learning.

A Closer Look: How Amorphic Data Cloud Enables Better Data-Driven Decisions

Cloudwick's Amorphic Data Cloud for AWS was at the core of data management and analysis:

1. **Amorphic Data Cloud Bulk Load Ingestion of Disparate sources to a Single Location:** Data from the Historian, MES and Vision systems were ingested into a single unified data store using **Amorphic Bulk Load Ingestion** feature improving time-to-value.
   - **Historian System Data**: Contains machine IOT data history captured from the Scada systems.
   - **MES System Data**: Data about the various work orders, product codes and execution schedules.
   - **Vision System Data**: Uses data from the vision scanner to detect reasons for product defects.

   The Amorphic Bulk Load connection automates the creation of connections with the source databases, extracting schema and creating relevant targets. The data is then loaded into the targets based on the schedules that can be set using time-based expressions. Amorphic's Bulk Load connection leverages scalable AWS services such as AWS Data Migration Service (DMS), Amazon S3, Amazon DynamoDB and many others to build production ready data pipelines in minutes.

2. **Data Centric Security with Role-Based Access Controls (RBAC):** Amorphic lets you optionally back up the ingested data from various sources into a data lake, which provides the ability for combined analytics. This process follows a data centric security approach over the data lake, which allows all datasets in Amorphic to be secured with role-based access controls (RBAC). Only users who have read or owner privileges can access and modify the data.

3. **Management of Structured and Unstructured Data at the DataSet Level for More Comprehensive Insights:** Structured MES data was ingested directly into the Amazon Redshift data warehouse for processing, whereas the Unstructured IOT data present in the Historian and Vision Systems was ingested into Amazon S3 making all data available in manageable formats for further aggregation and more comprehensive insights.

4. **Curated DataSets to Enhance Visualization:** Amorphic created low code ETL jobs designed to process the data within the Historian System and parse the data from the Vision System, making it ready for ingestion into Amazon Redshift. Redshift materialized views were used for further transformations and data aggregation to create curated datasets for visualization.

5. **OEE Dashboard Access in Real Time to Improve Data-Driven Decisions:** The OEE dashboards built on information from all three data systems – MES, Historian, and Vision - now curated and aggregated in the Amorphic Data Cloud assisted users in determining the optimal machine settings for different lines to drive higher performance.

6. **Machine Learning** experiments were also conducted on these curated datasets for OEE predictions. Insights gleaned from this analysis specifying which machine settings could drive higher OEE were shared throughout the organization for business users to access in real time and make improved data-based decisions.

Results

Ingesting all types of data – structured, semi-structured, unstructured – into the Amorphic Data Cloud for transformation removed potential data gaps, yielding more comprehensive data analytics that ultimately provided higher quality data recommendations. Amorphic removed the data silos by centralizing all data in one location and further provided the advanced analytics tools within the data cloud for the organization to run predictive analytics that business users could access in real time to make improved data-driven decisions. The visualizations we'll see in the next two paragraphs are examples of this output.

Correlation

Exhibit 8 below shows a sample data correlation between OEE, machine telemetry parameter Third Draw Pressure (bar) and production rate. A data analyst would highlight a target area of operation with OEE in range of 0.9, and Third Draw Pressure between 1.5-1.75 with a minimum color variation. Based on the data visualization, the operator will be advised to keep the production rate in the purple dots range of 583-729 for best outcomes.

Exhibit 8: Correlation Graph Between OEE & Third Draw Pressure

Machine Learning Recommendation

Exhibit 9 shows a Machine Learning Decision Path to reach a leaf node of OEE=0.97 in the decision tree. This corresponds to step 6 of Exhibit 6 i.e., building ML models (Decision Tree). The marked-up path recommends the settings of first draw pressure and second draw pressure to reach the leaf node of OEE=0.97.

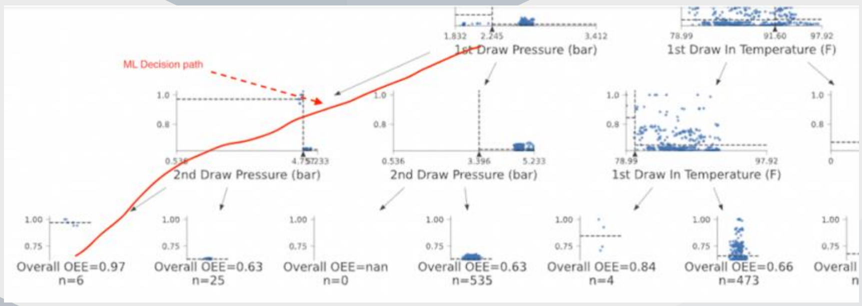

Exhibit 9: Machine Learning Decision Path

These recommendations provide data-driven guidance to the machine operators and supervisors to keep production metrics in the desired range.

Conclusion

Driven by rapidly developing technology and constantly developing markets, manufacturing companies must be agile with their data to get the highest quality insights to manage their operations and business goals. A best-practice data driven strategy should enable self-service answers from data, so users do not have to go to IT teams for every question or report to be run. Creating a strong data supply chain requires an agile data infrastructure which enables users to easily access data, share data and analyze data in real time with enterprise level security and governance in place. Organizations that have seen success with their Industry 4.0 initiatives invest in a production ready data cloud to simplify their data supply chain and create a future ready data driven ecosystem.

It is easy for manufacturing industry executives to be overwhelmed by IIoT, Industry 4.0 and Digital Transformation. However, as we've seen in this paper, implementing a data cloud can simplify the process of digital transformation by providing a single data repository to break down silos and provide a secure environment in which to obtain higher quality data insights on all data types. By using Industry 4.0 technologies, successful manufacturing organization can convert raw data into shareable insights, and use the insights to drive OEE, which translates into improved productivity, availability, and quality.

A production ready data cloud such as the Amorphic Data Cloud for AWS lowers the barrier to entry for Industry 4.0 digital transformation efforts, aiding manufacturing organizations in achieving higher operational efficiency as they embrace new business models. A Data Cloud such as Amorphic that centralizes and manages all data types in a single location while providing advanced analytics can be an essential first step in obtaining greater value from existing machine telemetry and MES data. Displaying OEE dashboards will elevate OEE manifolds. Centralizing data analysis for Machine Learning recommendations can provide incremental improvements in operational efficiency, and combined with complex data analysis, will drive competitive advantage and future-proof operations.