INTELLIGENT BUSINESS STRATEGIES

WHITE PAPER

Enabling Data-Driven Manufacturing using Data Virtualisation and AI

By Mike Ferguson
Intelligent Business Strategies
March 2024

Research sponsored by:

denodo

Table of Contents

INTRODUCTION – ASPIRING TO BECOME DATA-DRIVEN

Manufacturing companies are dealing with greater customer expectations, material shortages, inflation, and higher energy costs

Executives in most manufacturing companies today are facing major challenges in this post-pandemic economy. This includes dealing with greater customer and prospect expectations and also with things like geo-political tensions, material shortages, inflation, and higher energy costs, all of which can impact their supply chains and cost of manufacturing.

Buyers are demanding personalisation and to know about provenance of materials, and sustainability in manufacturing

With respect to customers and prospects, manufacturers are under pressure to provide them with more, to win their business and keep them loyal. For example, for business-to-consumer (B2C) manufacturers, offering customers and prospects the ability to configure products before purchase is now normal and expected. However, customers now want to know more from manufacturers to win their business. For example, they want to know more about the provenance of materials and about sustainability in the manufacturing process to produce the products they are looking to buy. Personalisation, fair trade supply chain practices and gaining the confidence of the “green buyer” all now matter.

In addition, smart manufacturing is high on the agenda, which requires a lot of investment in digital transformation including in robotics, IoT, cloud and AI. In this context, executives now recognise that data and AI have become mission critical in helping them run smarter.

Smart manufacturing is a priority and expectations as to the benefits of data and AI are huge

However, executive expectation of what data, business intelligence, machine learning and artificial intelligence (AI) can do for them is huge. Given the hype around AI, it is a very high bar. They see data and AI as the way in which they can offer personalisation, cut costs, improve efficiency, optimise manufacturing processes, and supply chains, automate tasks, improve quality control, reduce risk, and seize opportunity by making timelier, more accurate decisions. They also see it as the way to disrupt the markets they compete in and uncover new opportunities in new markets with new products and business models. In short, data and AI are now seen as central to the future of smart manufacturing.

Data and AI are becoming central to enabling smart manufacturing

A high quality, secure and compliant data foundation of business-ready data needs to be created to shorten time to value

For manufacturing companies to maximise their return on investment from data and AI, they need to build a high quality, secure and compliant data foundation upon which such competitive insights can be created, and AI exploited. It would be a grave oversight to just take this for granted. Data needs to be high quality, compliant, ‘business-ready’ and available to share across the enterprise (and beyond) to help shorten time to value and achieve the vision and strategic business goals that manufacturing executives are setting.

Data complexity is a major barrier to success

However, there are barriers in the way of success. One of the main challenges is data complexity caused by more and more data sources. Data now exists in data stores that are on-premises, in multiple clouds, and in software-as-a-service (SaaS) application provider data centres and at the edge. Many companies now have data persisted across a truly hybrid, multi-cloud distributed data estate, as shown in Figure 1. Data is stored in cloud and on-premises relational databases, NoSQL databases, cloud storage, file systems, and content management systems. It may also be streaming in from edge devices and ingested from multiple external data providers and suppliers via many different mechanisms such as APIs, file transfers, data marketplaces and data cleanrooms. So, we have a dilemma: on one hand, executive expectation is that data and AI provides the route to smart manufacturing, with more agility and

Growing IoT data volumes in multiple factories around the world together with the need to comply with data governance legislation in different geographies is adding to data complexity

speed, but on the other hand, the data needed to make it happen is becoming harder to find, govern and integrate. In addition, for global manufacturers with multiple manufacturing plants around the world, this is even more challenging particularly at the edge as IoT is deployed in multiple production lines in multiple factories around the world. In this case it is not just data complexity in terms of data sources, it is also about dealing with data volumes, scale and data governance legislation in different geographies that must be complied with, just to be able to continue to operate.

Many companies now have a hybrid, multi-cloud distributed data estate but still expect to make rapid use of data and AI to enable them to become more competitive

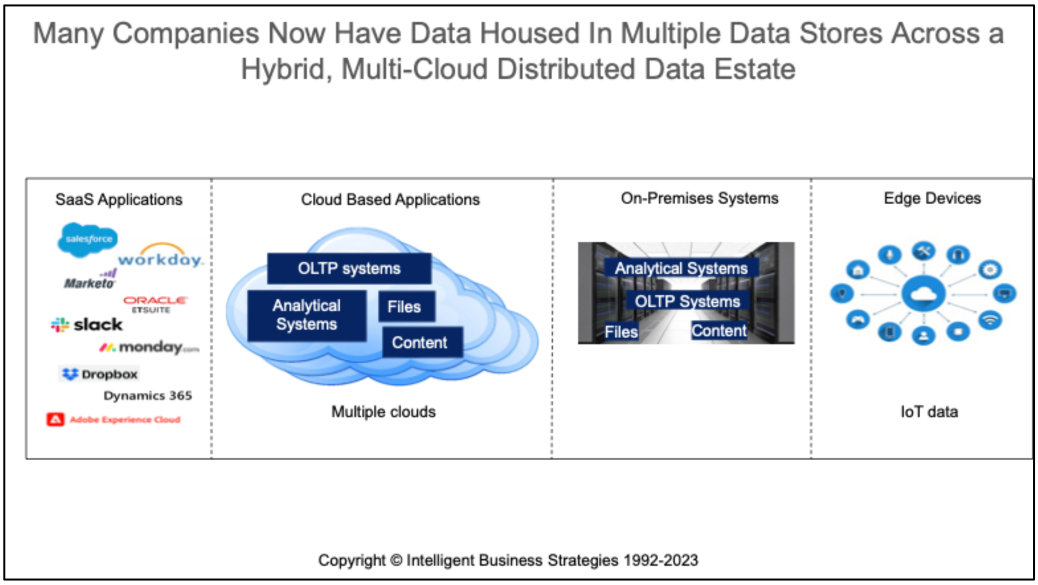

Figure 1

This paper looks at this problem and at how data virtualisation can provide a platform to help solve it so that manufacturing companies can shorten time to value, meet executive expectations and achieve their goal of becoming data-driven smart manufacturing organisations.

HIDING THE COMPLEXITY – THE NEED FOR DATA FABRIC

Given the data complexity that many manufacturing companies are facing, there are some critical technology components that need to be in place to speed up the creation of a data foundation that enables companies to build insights, predictions, recommendations, and AI-automation to help create a data-driven enterprise. This includes the need for data fabric software.

Data fabric software connects to data across a distributed estate enabling companies to find, engineer and govern data

Data fabric is data management platform software that is capable of connecting to the wide range of data stores and sources across a hybrid, multi-cloud distributed data estate. It offers the full capabilities to support automated data and relationship discovery, data classification, governance, cleansing, transformation, integration and sharing of data. It can be used by multiple IT professional and citizen data engineers to develop orchestrated pipelines that produce business-ready data products available for sharing securely across the enterprise in a compliant manner.

It often includes a data catalogue, support for collaborative development and a data marketplace

Data fabric software includes support for secure connectivity to a broad range of data sources, a data catalogue, collaborative development, support for DataOps, performance and scalability, extensibility, unified data governance and a data marketplace.

Its job is to hide the complexity, simplify data access and accelerate the development of data engineering pipelines to produce reusable data products that can be consumed to create machine learning models, AI bots, and a data warehouse to support business intelligence analysis and reporting. This is shown in Figure 2.

It hides the complexity of a distributed data estate

Multiple teams of data producers can use common data fabric to build pipelines that integrate data to produce reusable data products

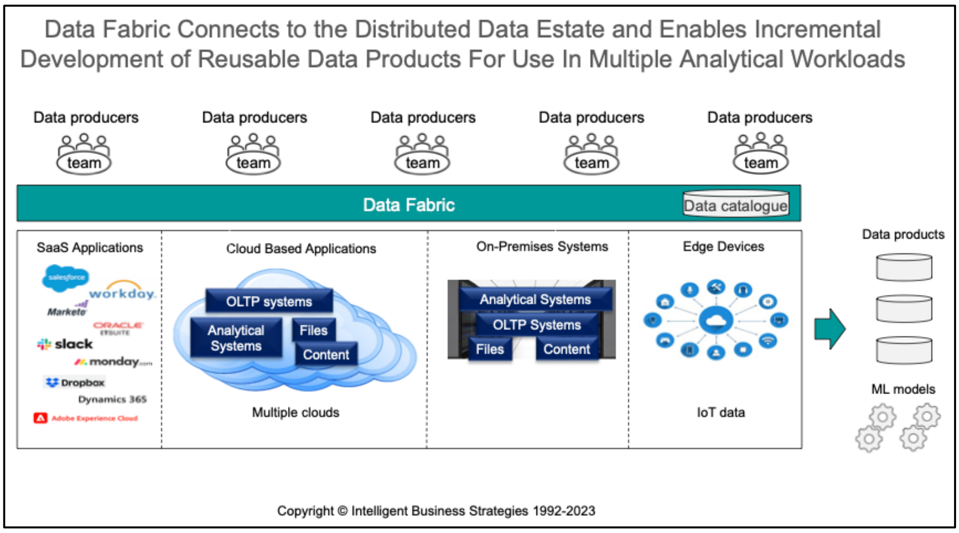

Figure 2

LOGICAL DATA MANAGEMENT FOR THE DATA DRIVEN ENTERPRISE

Logical data management software can be used as a data fabric platform to create integrated virtual data products from data in multiple underlying data stores.

Data products can be created by defining virtual views of data entities using common business data names defined in a business glossary. The platform's data catalogue provides business glossary capabilities together with metadata on source data needed to create each data product. Identified data in multiple data sources can then be mapped to the business terms and schema of each virtual view describing a data product. This approach allows you to *incrementally* create a layer of semantically linked virtual data products that can be:

Logical data management enables virtual data products to be created and consumed for different analytical use cases

- Rapidly assembled into different virtual schema optimised for different analytical workloads e.g., data warehouse, data marts, data science.
- Queried to provide feature data for machine learning model development in data science.
- Used as input to create other virtual data products.
- Consumed by other applications in manufacturing production lines via different interfaces e.g., SQL, REST, ODATA, GraphQL

This is shown in Figure 3.

Layers of virtual views can be defined that help to rapidly integrate data from multiple sources and make it available in easy-to-understand business terms for consumption

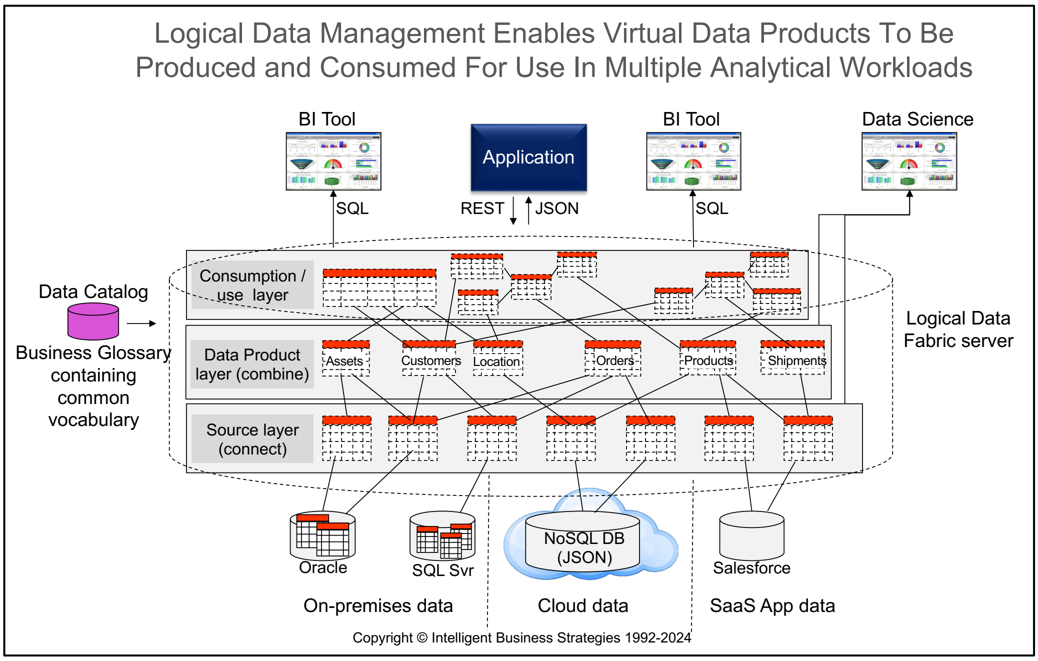

Figure 3

Data products can be quickly combined to create other virtual data structures

Data can also be persisted, if required

Note that this ability to quickly 'assemble' semantically linked data products to create other virtual schemas signals a major step change in productivity that increases agility and shortens time to value. It provides business-ready data to jumpstart multiple analytical projects as well as enabling component-based schema development. In addition, it enables a common semantic layer to be created so that all tools and applications are served data with the same common business data names.

MAJOR BUSINESS ISSUES IN MANUFACTURING

Manufacturing companies can benefit from the use of logical data management software

To understand the value of a data virtualisation and AI in manufacturing it is first important to understand current problems that manufacturing companies face. Then we can look at what the business priorities are and at the insights needed to help solve these problems. From there, we can look at the data sources needed and how the aforementioned logical data management can be used to integrate data from these sources to enable required insights to be easily produced.

PROBLEMS IN MANUFACTURING

Manufacturing companies are having to deal with rising production costs, labour shortages, material shortages, complexity and disruption of their supply chain and the need to improve quality control

There are several problems in manufacturing at present. They include:

- Rising production costs (raw materials, energy, labour)
- Labour shortage problems
- Material shortages, supply chain complexity and disruption
- Quality control in the product lifecycle and manufacturing process
- Digital transformation e.g. adoption of IoT, AI, automation
- Cybersecurity

Data and AI can be used to help reduce costs, improve efficiency, reduce waste, and increase automation

As in every industry, costs in recent years have been rising, whether that be cost of materials, energy costs or labour costs. Manufacturers often pass these costs on to the customer as price rises but they can also be proactive and look for ways to reduce costs by leveraging data and AI to improve efficiency, reduce waste and increase automation. Automation would also help deal with labour shortages. Beyond this, more can be done to minimise unplanned operational costs caused by equipment failure, product errors, shipment errors etc.

Data and AI can also be used to help monitor and optimise supply chains and production processes

Supply chain disruption and complexity also continues to plague manufacturing companies. Increasing supplier diversity will help but again, data and AI are becoming critical in helping to understand demand, supplier performance, and supply chain operations so that manufacturers can predict, act, and avoid disruption. Data and AI can be used to monitor and optimise supply chains, reduce the cost of raw materials, and improve materials planning. Predicting materials required from sales data helps optimise materials inventory, reduce lead times and avoid material transfers to overcome shortages in a specific manufacturing plant. Similarly, data and AI can be used to monitor energy costs, production throughput and minimise downtime.

Establishing a digital thread to provide consistent product and materials data to all systems involved in a product's design, manufacture and distribution improves product quality

Product quality is also a problem in production. One reason for this that bill-of-materials and product data needs to be available to all systems involved in the end-to-end production process of a product. However, frequently the way this has been done is to manually rekey this data into different systems repeatedly as the production process progresses. This can cause data entry errors that contribute to product quality errors. One way to help address this is to establish a digital thread from product design right through to product manufacture, distribution, and usage whereby integrated data can be shared and made available on-demand across the entire process.

TOP BUSINESS PRIORITIES IN MANUFACTURING

Manufacturing companies are embracing smart manufacturing offering product customisation, build-to-order, IoT for real-time data and AI automation

The entire process from front-office, to manufacturing, through to back office and the supply chain all need to be integrated

Integrated data for real-time visibility as well as BI, machine learning and AI are needed throughout the entire operation to truly embrace smart manufacturing

To overcome these problems, manufacturing companies are going through a major transition to embrace and adopt smart manufacturing practices. To do this involves integrating the business from end-to-end offering customers the ability to customise products to meet their personal needs and being capable of building-to-order. Also, if aspirations for smart manufacturing are to become reality, then more data is needed. It is not surprising therefore, that many manufacturing companies are investing in Internet of Things and AI to capture data to enable them to monitor and optimise their manufacturing processes, improve product quality and to establish resilience in their supply chain. But it is more than that. A manufacturing company needs to integrate the entire process from end-to-end by joining up front-office, manufacturing, and back-office processes as well as the supply chain. It also needs to provide integrated data, business intelligence, machine learning and AI-automation throughout the whole operation to truly embrace smart manufacturing. This is shown in Figure 1.

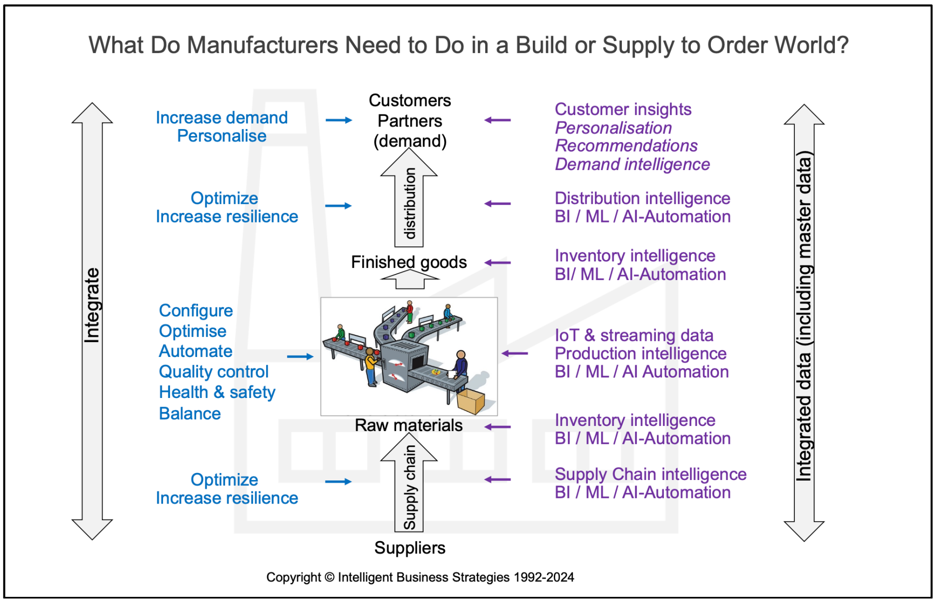

Figure 4

Given this requirement, manufacturing companies are currently prioritising the following areas:

Priority Area	Reasoning
Product customisation	• "Living customer profiles" • Be more competitive • Build-to-order • Buy before you build • Better customer service
Cost reduction	• Improved profitability, competitive advantage
Improving end-to-end efficiency	• Greater throughput and scalability with lower headcount
Smart factories	• More accurate demand forecasting • improve automation and equipment performance • AI-detected preventative maintenance

	• AI-assisted quality control to minimise product defects • Digital twins • Labour productivity tracking • Digital threads to track product performance and usage data from design through production, sale, use and disposal / recycle
Improving product quality	• Lower operational costs from not having to remake products • Less scrappage and waste • Improved customer satisfaction • Confidence in brand
Reducing waste	• Reduces costs, lower energy consumption
Supply chain resilience and optimisation	• Stronger materials planning • Better inventory management • Prevent disruption in the chain and of production • Avoid customer dissatisfaction caused by delays
Customisation / Personalisation	• Offer more personalisation allowing customers to tailor products to their specific preferences • Build-to-order model can improve cashflow
Regulatory compliance	• Health and safety • Product provenance • Customer confidence
Sustainability	• Utilising renewable energy sources • Optimising resource consumption • Recycling and compliance • Attract "green" buyers

Improving customer experience, smart factories, greater efficiency, better quality control, resilient supply chains, compliance, tackling cybersecurity, and sustainability are all priorities in manufacturing

Smart factories require automation, IoT, cloud computing and AI to reduce costs and improve efficiency

Looking at these priorities, the concept of a Smart Manufacturing is all about creating smart factories to reduce costs, improve efficiency and sustainability through the use of automation, Internet of Things (IoT), cloud computing, and artificial intelligence (AI). Other technologies like blockchain also help provide traceability and real-time visibility across the supply chain.

High quality, integrated, low latency data and intelligence is needed on suppliers and the supply chain to get products to market quicker

With respect to supply chains, manufacturers want a consistent, rapid approach to onboarding suppliers and better management of their suppliers to get products to market faster. The latter requires high quality, enriched, integrated supplier data and detailed supplier intelligence. In addition, integrated data yields greater supply chain transparency. Manufacturers also want to use AI and lower latency data to improve supply chain agility, sustainability, and resilience.

Rea-time monitoring of events is also needed to detect, predict, recommend and provide timely actions to avoid disruption, prevent waste, maximise yield, optimise inventory and improve quality control

All of this requires the ability to access and share data across the supply chain. It also requires monitoring of events in real-time to detect and predict problems and act in a timely manner to avoid disruption, prevent waste, complete work orders, and continue to deliver on-time before these events and problems have a major business impact. Sharing integrated data also helps certify product materials and the origin of products thereby providing product provenance and compliance. It also allows manufacturers to optimise inventory and demonstrate that they have control of product quality to their customers. In addition, in the event of a product recall, detailed data across the supply chain enables disputes to be resolved quickly and at lower cost.

INSIGHTS NEEDED TO SOLVE MANUFACTURING PROBLEMS

Integrated data is needed to produce richer, more comprehensive insights to support better decision making to help address problems

To solve these problems and meet their top business priorities, manufacturing companies need to integrate and utilise data to produce richer and more comprehensive insights that support decision making in each or the aforementioned high priority areas. That means integrating data from a range of different internal and external data sources to create insights associated with the areas shown in the table below.

Insights are needed on the market, demand forecasts, production efficiency, production costs, product quality, inventory levels, suppliers, supply chain and sustainability

Type of Intelligence	Insights needed
Market intelligence	• Single customer view • Customer segmentation • Personalisation • Marketing automation
Demand planning and forecasting	• Sales forecasts
Production efficiency intelligence	• Production schedules • Equipment performance levels • Production rates • Work order completions • Bottlenecks • Manufacturing events (e.g. equipment failures)
Production cost intelligence	• Material costs, energy consumption levels and costs, labour costs in manufacturing processes
Production quality intelligence	• Scrap rates, defect rates, patterns, trends, root cause analysis
Inventory levels and predictive inventory management	• Raw materials • Finished goods
Supplier intelligence	• Supplier locations • Supplier relationships • Materials provided • Supplier delivery performance • Supplier lead times • Supplier risk and compliance • Supplier sustainability • Total suppliers spend to enable better negotiation of discounts, pricing, and payment terms
Supply chain intelligence	• Materials inventory by plant • Delivery shortages • Internal events (e.g., new orders, changed orders, cancelled orders) • External events (e.g., strikes, weather, congestion)
Sustainability	• ESG reporting

Machine Learning Models and AI Needed in Manufacturing

Machine learning and AI automation are also needed

Both require high quality and timely integrated data

In addition to insights from business intelligence (BI), manufacturers also need to make use of machine learning and AI to help meet top business priorities.

Much like BI, both machine learning and AI also require integrated data to build the necessary predictive and prescriptive models to detect and predict events, identify problems and opportunities, and to make recommendations and/or take automated actions to help optimise supply chain and manufacturing operations, reduce risk, improve inventory levels and more.

There are many areas in manufacturing that need to make use of machine learning (ML) and AI. Examples of the types of models needed are shown below.

Machine learning is needed to prevent equipment failure, predict inventory, improve quality control, maximise yield, improve health and safety, optimise the supply chain, and implement cybersecurity

Model Type	Business Purpose
Predictive and preventative maintenance models	To avoid unplanned downtime due to equipment failure
Predictive inventory models	To avoid shortages and ensure customers get high quality products on time
Quality control image recognition models	To control product quality, detect quality issues in the production process, ensure corrective action is taken and avoid delays
Prescriptive equipment optimisation models	To ensure equipment performance is running optimally to avoid delays and maximise yield
Prescriptive risk models for health and safety risk reduction	To predict, detect and act on risk issues to ensure health and safety at work
Prescriptive models for dynamic planning, forecasting, and scheduling	To continually keep the end-to-end process optimised to meet demand while also enabling the business to scale
External event warning models for supply chain optimisation	To monitor, predict and act on external events to avoid disruption in the supply chain and manufacturing process
Cybersecurity detection models	To detect and prevent unplanned outages and business disruption caused by cyber security incidents

Predictive and preventative maintenance models can be trained on asset data and IoT data like temperature, pressure, and vibration readings from the machines used in the production process.

Quality control image recognition models taking live data from cameras deployed at particular stages on production lines, can act as scrap advisors and recommend equipment parameters to machine operators to reduce waste.

Besides the traditional machine learning examples shown, it is also possible to make use of generative AI in manufacturing. Although it is still early days, generative AI can be used to:

- Generate design alternatives in the product design process
- Generate code to program machines and so help to reduce the time and effort involved in introducing automation

Generative AI can be used to improve productivity in design, programming equipment, preventative maintenance, and material handling

- Generate text or images to provide step-by-step instructions for preventative maintenance engineers to shorten maintenance times, ensure correct maintenance procedures are followed and get the most out of equipment. This could include a list of any replacement parts needed
- Translate equipment operator natural language prompts into a sequence of actions to perform material handling tasks on a piece of equipment like a robot
- Generate product information and marketing content more rapidly to help bring products to market quicker

Traditional machine learning and generative AI are critically important to the implementation of smart factories and smart manufacturing. The development and subsequent retraining of all of these models requires data from multiple data sources to be captured, cleaned, transformed, and integrated which means data engineering is central to success. This is true not just for initial model development but also to model re-training to ensure model accuracy is maintained over time.

AN INCREASING NUMBER OF REQUIRED DATA SOURCES

The number of data sources now needed by manufacturing companies to provide data for analysis is increasing rapidly

The use of business intelligence, machine learning and AI in manufacturing is fundamentally dependent on capturing and integrating data from a number of sources across the supply chain and also across the end-to-end manufacturing operational process.

Access to data is needed to provide:

- A 360 view of suppliers, materials, products, assets, and customers
- Data for data warehouses and data marts to enable business analysts to produce the different kinds of business intelligence discussed earlier
- Data for model development and building generative AI bots in data science
- Real-time event stream processing and streaming analytics to monitor supply chain operations, production line operations, and customer events (e.g. changed orders, cancelled orders) as well as product performance and usage once sold

The data sources needed to provide data to support these analyses includes:

- Customer Relationship Management system data
- ERP system data
    E.g., production plans, material requirement plans, shop floor activities, work orders, production schedules, and inventory data
- Procurement / supply chain management system data
- Programmable logic controllers (PLCs)
- Manufacturing execution systems (MES)
- Supervisory Control and Data Acquisition (SCADA) systems
- Quality control systems
- Vision systems
- Industrial IoT data (streaming off production line equipment) typically in MQTT format
- Energy monitoring systems
- Asset management systems
- Product-lifecycle management systems
- Inventory management systems
- Distribution systems (for shipments)
- Weather data
- External supplier data
- Master data (customers, products, suppliers, assets etc.)

The data required to support analytical workloads comes from a range of internal and external sources and devices scattered across a hybrid, multi-cloud distributed data estate

Note also, that in many cases, more and more of the transaction applications in use are SaaS applications. Therefore, data needed can be scattered across multiple SaaS applications, on-premises data stores, different data stores in multiple different clouds as well as a huge amount of live streaming data coming

Data needs to be integrated for analysis and also to provide 360º views of customers, products, suppliers, materials, plants, and assets

off IoT devices and clickstream data. It is exactly like the distributed data estate shown in Figure 1.

Data from these systems needs to be integrated not just for analysis but also to create 360º views of customer, product, supplier, materials, plant, and asset master data to drive consistency across systems. Other key data entities that are needed include customer orders, work orders, raw materials and finished goods inventory data and shipments. However, it is not just about data integration. To enable confidence in data, there is also a need for a business glossary of common business data names (a common business vocabulary), data quality and lineage.

A business glossary of common business data names is needed to share data and get visibility across a supply chain

A business glossary of common business data names is critical to getting visibility across the supply chain because data is being supplied by multiple organisations with many different physical data names for the same data. Common business data names help to ensure everyone is clear on the meaning of the data being shared. The same is true if manufacturers are to create digital threads to share data across systems in the end-to-end design and manufacturing process.

Also, with respect to IoT data, this data may need to be accessed in real time and integrated with other data (e.g., master data) to enable monitoring of live production and supply chain operations. Figure 5 shows the contrast between real-time IoT data processing versus data integration for a classic data warehouse. Both are needed in smart manufacturing.

Integration of real-time IoT data with data on assets, products, materials, and suppliers is also needed to continuously monitor the supply chain, manufacturing, and distribution

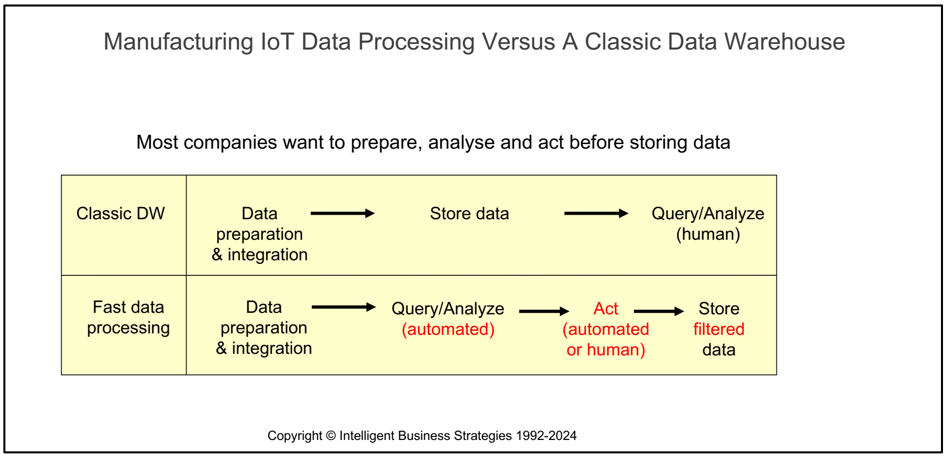

Figure 5

Using Data Virtualisation in Manufacturing

The major challenge to overcome is the ability to engineer and integrate data quickly

Looking at the insights needed and the growing number of data sources that manufacturing companies need data from, one of the most significant challenges that has to be overcome is the ability to engineer and integrate data quickly.

On-demand data integration is needed in every area, whether it be customer 360° live profiles, supply chain, the manufacturing process, inventory management, distribution, risk management and compliance.

Data fabric software with data virtualisation capability enables on-demand data integration

Data fabric software can connect to the broad range of required data sources to integrate data needed by specific kinds of analyses. This is where the data virtualisation capabilities of logical data management play a vital role. These capabilities help to integrate the data needed to provide insights in all of the manufacturing use cases mentioned. Let's look at three key examples.

Smart Manufacturing

The purpose of smart factories is to prevent unplanned equipment failure, automate tasks, lower costs, improve product quality, prevent errors, and optimise manufacturing processes to improve efficiency and lower costs. To create a smart factory requires business intelligence, machine learning and AI.

Data needs to be integrated to provide consistent data on-demand across the manufacturing process and to support multiple analytical workloads

That means data needs to be captured, cleaned, and integrated to enable:

- Consistent integrated data to be available on-demand to all systems involved in the manufacturing process as it executes to complete a work-order e.g. a digital thread
- Historical data to be built up and made available for analysis to produce intelligence associated with the entire end-to-end manufacturing process as shown in Figure 4. This includes creating intelligence on production efficiency, production costs and production quality
- Data to be made available to build machine learning models and AI bots
- Data to be made available to deployed ML models and AI bots
- Real-time monitoring of live operations with respect to monitoring equipment performance, efficiency, costs, and quality control
- Data to be made available to construct a digital twin for use in simulation and analysis

Figure 6 shows how data from disparate data sources can be integrated using logical data management data virtualisation capabilities to help enable smart manufacturing. It can be done using a layered approach in Figure 3 that includes the creation of data products or as a logical data warehouse integrating multiple analytical systems or a combination of both.

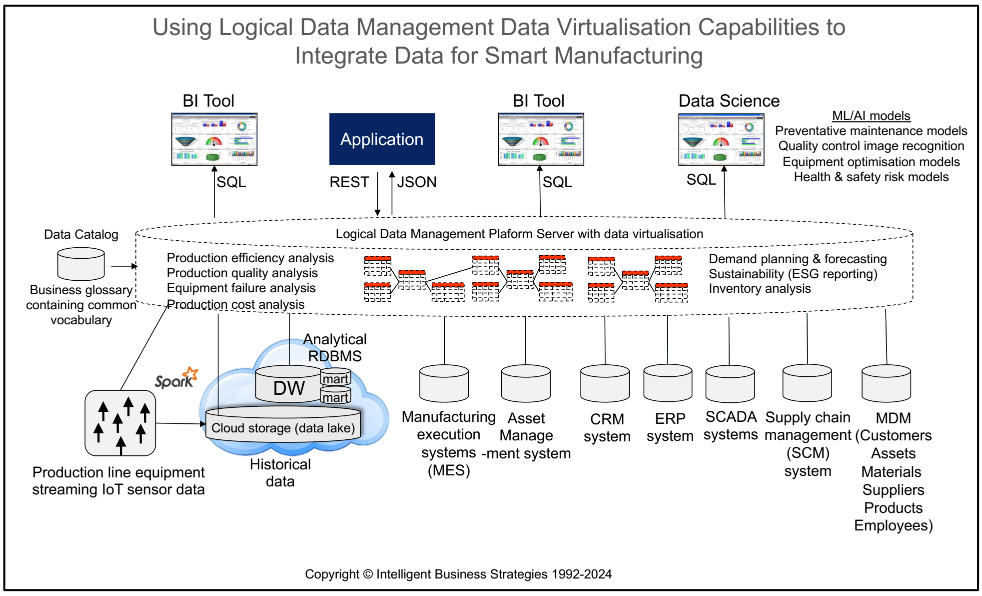

Figure 6

Data can be accessed on-demand about sales, work orders, material deliveries, inventory, machinery, and how it is performing during the manufacturing process as well as historical data on costs, efficiency, and yield

Here you can see how data can be accessed about:

- Customer demand and orders
- Which products to manufacture, work orders, material delivery, inventory and use, and shipping all from the ERP system
- Manufacturing assets, e.g. equipment, parts, production lines etc.
- How products are manufactured, e.g. data on operations scheduling, production activity and resource allocation from MES systems
- Industrial equipment performance (streaming IoT data)
- Controlled machinery logs in SCADA systems
- Historical data on production performance and production quality (Data Warehouse / data lakehouse)
- Customers, products, suppliers, materials, assets etc.,

Master data helps provide context

Existing analytical systems may already exist, such as a data warehouse

Smart Manufacturing Insights

Multiple virtual views can provide data for end-to-end business intelligence, machine learning and AI

Data from operational and analytical data sources can be integrated and virtual views created to provide holistic and near-real-time views to support different analyses such as end-to-end intelligence on marketing, production efficiency, production quality, inventory, and shipping.

Digital Threads

Data can be integrated and shared in a digital thread to prevent unplanned errors, delays, and costs in production

Another critical requirement is improving data quality to prevent unplanned errors, delays, and costs in production. This includes getting consistent product data across multiple systems from product design all the way through to production. This is referred to as a digital thread. Different integrated views of this data need to be available in real time for query by systems and people at different points in the manufacturing process during manufacture.

Data that needs to be shared in a digital thread may include:

- Product data (including bill of materials)

Consistent virtual views of integrated data can be made available in a digital thread using a common business vocabulary

- The origin of materials and parts / components used
- Engineering changes, product variants, substituted materials, etc.
- Work orders, machine operator instruction data, and IoT data associated with a work-order or production lot
- Process stage test and inspection data, e.g., pass / fail results, corrective actions, etc.

This also facilitates real-time visibility and operational reporting

Note that if digital threads are to be successful, a common business vocabulary needs to be created, as many different data sources will have different physical data names for the same data. Without common business data names, it is nearly impossible to provision consistent data to all systems that need it. Also, by providing access to this data via integrate virtual views, it is possible to monitor demand, work-orders, production plans and inventory levels in real-time for operational reporting from one end of the manufacturing process to the other.

Digital Twins

It is also possible to provide data to create different kinds of digital twins to help identify design flaws early, reduce waste, improve yield and sustainability, and reduce time to market

As more of this data is captured, it can be integrated with IoT data and used to build digital twins to help analyse and optimise the performance of individual pieces of complex equipment in the production process and to help optimise the manufacturing process as a whole. Different kinds of digital twins can be created, and the benefits can be significant including:

- Reducing time to market
- Early identification of product design flaws
- Reduction in scrap waste
- Improve manufacturing yield
- Sustainability improvements

Industrial IOT

At the heart of every smart factory is the use of industrial IoT sensors and devices that can be embedded within machinery and deployed throughout the entire factory. These sensors and devices continuously emit data in real time on humidity, temperature, pressure, vibration, etc., to continuously monitor things like equipment status, equipment performance, environmental conditions (e.g., energy usage), health and safety risks etc. Raw IoT data is needed to ensure equipment is running optimally at maximum capacity to avoid delays and maximise throughput and also to ensure costs are minimised and employee health and safety is not compromised. Product quality can also be monitored.

IoT data can be accessed and combined with asset data to provide real-time visibility of manufacturing operations for use in automated analysis and situational awareness

Most of this needs to happen close to where the data originates simply because of the sheer volume of raw IoT data that can be generated. However, IoT data often needs to be integrated with asset and materials data at run time to provide context that makes real-time predictive and prescriptive analysis possible. Also, making raw IoT message data available using common business data names and using the data types that don't round down IoT readings is also important.

Data virtualisation can connect to IoT streaming data sources like Kafka and to data at rest to provide integrated virtual views of real-time manufacturing data using common business data names. This enables the creation of situational awareness dashboards and machine learning models to be used for automatic analysis to:

This can be used to predict equipment failure, bottlenecks, health and safety risks and quality issues so that actions can be taken in a timely manner to avoid problems, protect employees and keep everything optimised

- Predict equipment failure and schedule preventative maintenance to avoid unplanned downtime
- Detect bottlenecks and recommend adjustments to optimise equipment performance
- Detect quality issues in the production process and ensure corrective action is taken so that scrappage and delays can be avoided
- To predict, detect risk and act to ensure health and safety at work

COMPLIANCE AND RISK

In a global economy, many manufacturers not only sell products in many different countries, but they may also have manufacturing plants in many different countries that make these products. Global manufacturers therefore have to comply with an increasing number of country specific and regionally specific regulations and legislation covering a wide range of obligations. This includes regulations and legislation associated with:

Manufacturers have to comply with an increasing number of regulations and legislation covering a wide range of obligations

- Product safety
- Health and safety
- Environmental impact
- Data handling and protection
- IT safety and security
- Fair trade and competition
- Export controls
- Anti-corruption laws
- Employment laws

Regulations are particularly strict around things like the manufacture of pharmaceutical drugs, medical devices, solar panels etc. This includes environmental safety. For example, in the European Union, the European Agency for Safety and Heath at Work has for many years had a European Machinery Directive that applies to the manufacture of machinery, interchangeable equipment, safety components, lifting accessories, chains, ropes and webbing, removable mechanical transmission devices and partly completed machinery (e.g., a drive system) intended to be incorporated into or assembled with other machinery.

The regulatory burden across multiple legal jurisdictions is significant and requires data from many different data sources

Needless to say, the challenge of dealing with this regulatory burden across multiple legal jurisdictions is significant in that it requires data from many different data sources (which in many cases are in different geographies) to be integrated to enable regulatory reporting and auditing. Traceability and transparency are also paramount.

Data virtualisation can be used to provide integrated virtual views of data needed for compliance reporting and monitoring risk

To overcome this problem, data virtualisation can be used to create integrated virtual views of real-time data to provision the data needed for compliance reporting and monitoring risk. Different virtual views can be created to enable reporting in different legal jurisdictions without the need to copy data. This hides the complexity of data integration from business users allowing them to deliver compliance reports quickly and in a timely manner.

SUPPLY CHAIN MANAGEMENT

Manufacturers want real-time visibility across the supply chain and also resilience to avoid disruption

With respect to supply chain the challenge for manufacturers is to have real-time visibility across the chain and build resilience into it. To do this, manufacturers need to future-proof supply chains by:

- Investing in advanced technologies like blockchain for enhanced transparency and traceability

Diversifying sourcing is important

- Diversifying sourcing and investing in local production to establish multiple suppliers of the same materials servicing each manufacturing plant

Creating common business data names is critical to enable data sharing

- Creating a common vocabulary of agreed-upon business-friendly data names to understand and share data across the supply chain
- Connecting to multiple internal and external sources and creating common business vocabulary virtual views of real-time data across the supply chain (without the need to copy or consolidate data) for real-time supply chain visibility

Data virtualisation can be used to create integrated virtual views of data to help provide intelligence on suppliers and real-time visibility of supply chain events and operations

- Creating integrated virtual data products that allow them to build a comprehensive set of:
  - Up-to-date supplier intelligence including understanding supplier locations and proximity to manufacturing plants, materials provided, alternative suppliers, supplier delivery performance, total supplier spend, supplier lead times, supplier risk and compliance etc.
  - Real-time supply chain intelligence including real-time views of materials inventory by plant, deliveries and delivery shortages, customer orders and changes to orders, delivery vehicle locations, external events etc.

Real-time virtual views of the supply chain enable integrated data to be provided on-demand for situational awareness dashboards and for automated monitoring using AI

- Accessing real-time virtual views to create always up-to-date operational situational awareness reports and dashboards to allow each manufacturing plant to see its supply chain
- Building and using machine learning models to automatically monitor the supply chain in real time:
  - To automatically detect and predict disruption
  - Using predictions together with supplier intelligence and real-time supply chain intelligence to cause automated alerts, recommendations, and actions to prevent disruption and keep the supply chain optimised

This helps to avoid disruption and keep the supply chain optimised

  - To automatically act to avoid external events disrupting the supply chain and the manufacturing process

Figure 7 shows how data from disparate data sources across the supply chain together with an internal data warehouse can be integrated using logical data management data virtualisation capabilities. This helps to create real-time supply chain visibility and provide the data needed to monitor and make timely decisions across the supply chain to keep it optimised. This includes providing real-time situational awareness reports.

Real-time visibility of the supply chain enables changes in demand and problems to be detected early or predicted and actions taken to avoid disruption

Supplier performance and spend can also be monitored

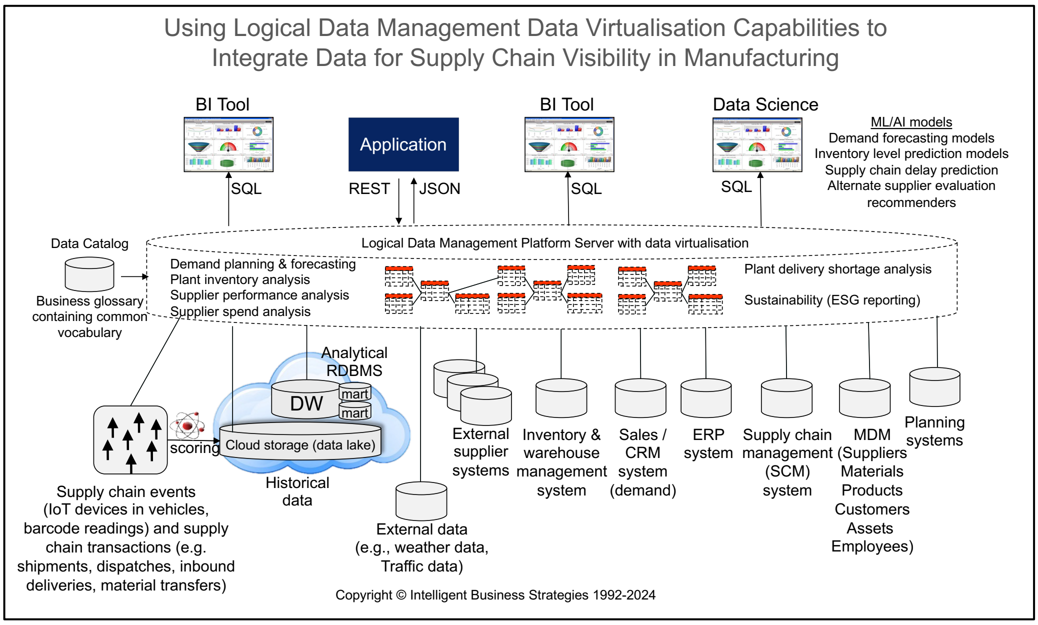

Figure 7

Enabling a Supply Chain Control Tower for Visibility

A supply chain control tower can be created offering different personalised views of the chain for end-to-end situational awareness

Also, by integrating data from internal and external data sources, it becomes possible to create a supply chain control tower consisting of a connected set of personalised dashboards aimed at different parties in the supply chain. This includes your suppliers, your organisation as a manufacturer, your distributors, and customers. These different personalised views can display real-time metrics and events across the supply chain for end-to-end situational awareness. This could also be combined with a digital twin of your own manufacturing plants and your customer's network and live GPS data to show status at different manufacturing plants and customer locations.

Supply Chain Command Centre for Pro-active Optimisation

A supply chain command centre includes a control tower but also event monitoring, alerts, recommendations, task, automation, task assignment and orchestration to keep the chain optimised

It is also possible to create a supply chain command centre using logical data management software. This goes beyond a supply chain control tower by providing data integration, supply chain visibility, work orchestration, recommendations, automated task assignment to operators, and task automation to enable resilience and pro-active end-to-end supply chain optimisation.

For example, access to integrated data by prescriptive machine learning models makes it possible to offer live alerts and recommendations across the supply chain, automatically assign tasks and trigger automation to help all parties keep the chain optimised. The result is an integrated sense-and-respond capability.

Virtual data views can also be created to support customer 360^{o} marketing, sales, and service, cybersecurity, and more

In addition to these examples, many other kinds of analyses can be undertaken in manufacturing by discovering and integrating data using logical data management. This includes marketing, sales, and customer service oriented analyses using data in a customer 360^{o} data platform, and cybersecurity analyses.

Conclusions

Manufacturing companies must be able to integrate data from multiple data sources in a timely manner if they are to become data and AI driven enterprises

If manufacturing companies are to become data and AI driven, their future is heavily dependent on their ability to integrate data from an increasing number of data sources in a timely manner. This requirement is now becoming mission critical to building a high quality, secure, compliant, and rich data foundation to help produce new insights, recommendations, and predictions. It is this foundation that will enable them to create the insights they need to provide visibility of the supply chain, the end-to-end manufacturing process, digital threads, digital twins, and monitoring of live production line operations.

Logical data management hides the complexity of a distributed data estate, improves agility, and can expedite data integration

The main challenge is overcoming the complexity caused by an ever-increasing number of data sources and data stores scattered across a truly hybrid, multi-cloud, distributed data estate. Implementing logical data management software that can connect to these data sources and enables the building of virtual data products while allowing source data to stay where it is, will help improve agility, overcome bottlenecks in data engineering and incrementally deliver the insights needed in a timely manner to improve business outcomes.

Multiple teams of data producers can create reusable virtual data views that shorten time to value

Logical data management offers these capabilities and is therefore a candidate technology that manufacturing companies should shortlist when selecting data management technology for smart manufacturing and a data-driven enterprise.

About Intelligent Business Strategies

Intelligent Business Strategies is an independent research, education, and consulting company whose goal is to help companies understand and exploit new developments in business intelligence, machine learning, advanced analytics, data management, big data, and enterprise business integration. Together, these technologies help an organisation become an *intelligent business*.

Author

Mike Ferguson is Managing Director of Intelligent Business Strategies Limited. As an independent IT industry analyst and consultant, he specialises in BI / analytics and data management. With over 40 years of IT experience, Mike has consulted for dozens of companies on BI/Analytics, data strategy, technology selection, enterprise architecture, and data management. Mike is also conference chairman of Big Data LDN, the largest data and analytics conference in Europe and a member of the EDM Council CDMC Executive Advisory Board. He has spoken at events all over the world and written numerous articles. Formerly he was a principal and co-founder of Codd and Date Europe Limited – the inventors of the Relational Model, a Chief Architect at Teradata on the Teradata DBMS, and European Managing Director of Database Associates. He teaches popular master classes in Data Strategy, Data Architecture Modernisation, How to Govern a Distributed Data Estate, Practical Guidelines for Implementing a Data Mesh, Data Catalogs and Embedded Analytics, Intelligent Apps and AI Automation.

Telephone: (+44)1625 520700
Internet URL: www.intelligentbusiness.biz
E-Mail: info@intelligentbusiness.biz