#25-052
2025 AFS Proceedings ©American Foundry Society

The 6C Framework to Build a Connected Factory

David J. Blondheim, Jr.
Mercury Marine, Division of Brunswick, Inc., Fond du Lac, Wisconsin, USA

ABSTRACT

Embracing Industry 4.0 (I4.0) relies on the crucial role of data collection and utilization. The benefits of I4.0 drive operational excellence in connected factories by elevating productivity, uptime, and quality. While the rationale of Industry 4.0 adoption is widely acknowledged, the challenge lies in the practical implementation of a connected factory. The resolution of both technical and human challenges is required for successful adoption. The proposed 6C Framework is structured around six key components to solve these challenges: *Criteria*, *Connect*, *Communicate*, *Collect*, *Consume*, and *Culture*.

Understanding this high-level framework is foundational, as it guides countless strategic management decisions when implementing data collection. This framework is deconstructed into tactical aspects to ensure proper technical and cultural questions are considered throughout the data collection process. The 6C Framework addresses the human and technical aspects of data collection to aid in implementing I4.0.

Keywords: Industry 4.0, I4.0, smart manufacturing, internet of things, IoT, operational technology, OT, information technology, IT, data collection, database, framework, culture, data

INTRODUCTION

Manufacturing is experiencing a significant and difficult change. The rapid revolution of technology conflicts with the evolutionary pace of adaptation and implementation. This dichotomy between the revolution and evolution is unfolding in the digital transformations occurring now in manufacturing. This new transformational era is called the Fourth Industrial Revolution, or to use the term coined in 2011, Industry 4.0 (I4.0).[1,2]

Just as previous industrial revolutions disrupted manufacturing, Industry 4.0 brings significant change to factories. The change is especially challenging because I4.0 is transforming all aspects of the factory while building on top of the previous industrial revolutions. The interchangeable parts that created effortless assembly in the first industrial revolution are now the communication protocols transferring data seamlessly between distinct software systems. The massive volumes of data generated by manufacturing processes have succeeded in the mass production of assembly lines since the second industrial revolution. Industry 4.0 supercharges and connects the computers, robots, and controllers brought to manufacturing in the third industrial revolution. The creation of these complex cyber-physical systems is what defines and distinguishes the fourth industrial revolution from the third.

Integrating technologies, digital data, and the people involved is the challenge manufacturers must overcome today. The first two industrial revolutions combined new inventions with the people operating these processes. The third industrial revolution introduced a triad, adding the digital data generated from these systems. Much like Henry Ford's assembly lines amplified the concept of interchangeable parts, the cyber-physical network of I4.0 accelerated and expanded the computer, robotic, and controller technologies developed in the previous industrial revolution. Manufacturing is digitally transforming, and companies must adapt.

I4.0 IS DIFFICULT TO EXECUTE

Industry 4.0 faces a dilemma in the speed dichotomy mentioned earlier. Research has highlighted that the first decade of implementation of I4.0 has been poor. A McKinsey study showed 6 out of 10 manufacturers faced barriers so strong they achieved limited to no progress on their I4.0 initiatives.[3] In 2018, a study of German industry showed 9% of companies were able to implement a comprehensive Industry 4.0 approach within their organization.[4] In 2019, this dropped to 8%[5] and by 2022 it was 7%.[6] Numerous barriers have been identified as reasons for this low level of implementation: financial issues, organizational challenges, IT infrastructure maturity, company size, lack of employee skill sets, risk of security breaches, disruption to existing jobs, and general resistance to change.[1,7–9] The lack of resources and employees' competency and expertise are regularly identified as the most significant barriers for adopting I4.0.[9]

Many barriers that manufacturers face today with Industry 4.0 parallel those encountered during the development of the assembly line in the second industrial revolution.

Henry Ford pioneered the use of the assembly line to revolutionize automobile manufacturing in the early 1900s. By leveraging interchangeable parts and incorporating new technology, Ford upended the automobile industry. This transformation did not come easy. After much trial and error to perfect the assembly line in 1913, Ford immediately faced a worker shortage as employees left for competitors, complaining the work was now too simple and boring.[21] Ford solved the technology problem, only to immediately be confronted with an employee culture problem. Ford's success came from addressing both issues.

Today's manufacturers will achieve the same success by merging the triad of the challenges they face: technology, data, and employees. To move forward, manufacturing management must understand how these three pieces fit together. A framework is needed to guide the connecting and communicating with equipment, collecting data, and providing it to employees to use. The human aspects of I4.Management must fully appreciate 0. Following this technical and human framework will foster a data-driven culture that embraces the use of data in manufacturing.

The digitization of manufacturing brings programming and software front-and-center for factories, which historically are slow to adopt new technologies. Advances in areas like artificial intelligence (AI), machine learning (ML), and cyber security are difficult to fully understand, and change at blistering speeds. Applying new technology which constantly changes creates hurdles for companies to find the right skill sets, train employees, and dedicate the financial resources required for I4.0. Manufacturing historically is centered around durable capital equipment. Installation of this type of hardware moves at a snail's pace compared to improvements in AI.

Manufacturers in each of the industrial revolutions have faced this same dilemma. How do companies successfully implement new technologies and embed the culture needed to be successful with change? A roadmap or framework is needed to help guide manufacturing management.

I4.0 IS DIFFICULT TO DEFINE

One challenge for manufacturers is to define all the technologies encompassed in I4.0. Implementing I4.0 can mean different things to different companies. For one company, I4.0 could mean implementing virtual reality to help train maintenance employees. Another factory might implement the latest networked industrial robots to speed up production, while a third could utilize simulation software to predict outcomes in production processes. Some technologies, like blockchain, rose and fell from prominence over a few short years but are still cited as I4.0 technologies. Others, like artificial intelligence (AI) and machine learning (ML) continue to rapidly evolve with the release of large language models and other chat-bot applications. Additive manufacturing is often stated as a key I4.0 technology, which contrasts with pillars like cybersecurity, data analytics, and digital twins.

There are two key points to consider regarding I4.0 technologies. First, there is no universal definition associated with the number of technology pillars within Industry 4.0. Authors have their own opinion on groupings, numbers, and importance. Multiple publications state there are anywhere from four to 14 different I4.0 technology pillars.[10–14] Second, the technology pillars are extremely diverse. Some of the technological pillars of I4.0 are not compatible with all companies. For example, a company does not have to implement blockchain or additive manufacturing (AM) to have successfully implemented I4.0, especially if these technologies do not add value to the organization. The diversity of these technologies makes it confusing to discuss the process of implementing I4.0 in industry, as the term can represent many different technologies.

I4.0'S BENEFITS CAN BE ACHIEVED

The benefits of I4.0 must outweigh the challenges associated with the implementation, given the industry's drive to adopt these technologies. The author believes the revolution brought by I4.0 lies in fully integrating the equipment, data, and people involved in manufacturing to achieve a step-change improvement in productivity, uptime, and quality.[15–17] This view is supported by literature highlighting the benefits of Industry 4.0.[18–20] Ultimately, operational excellence will be achieved through the utilization of real-time, data-driven decision-making across all aspects of manufacturing. While some decisions will be automated through computers and machine learning, a significant portion of decisions, especially in the near term, will continue to rely on employees.

To narrow the scope within this publication, the I4.0 technologies focused here relate to data connection, collection, and consumption in building a connected, or smart factory. Gaining the benefits of I4.0 occurs only when the company's employees are aligned with the technology being implemented. This paper introduces the 6C Framework for Building a Connected Factory. It emphasizes six unique phases:

Criteria, *Connect*, *Communicate*, *Collect, Consume* and *Culture*.

The framework aims to assist manufacturing management in addressing both technical and human challenges associated with implementing Industry 4.0 technologies in a connected factory.

6C FRAMEWORK

The 6C Framework presented in this paper is derived from the author's experience in implementing data collection, data analysis, and I4.0 technologies across various manufacturing factories and processes. Although other I4.0 frameworks exist,[12,22–24] they are often highly strategic, conceptual, or academic. These frameworks lack the tactical details or practical roadmap that small- and medium-sized manufacturing facility can immediately use to implement I4.0. The 6C Framework aims to address these shortcomings.

The 6C Framework not only structures data collection and consumption in manufacturing factories but also serves as a roadmap. The 6C Framework starts with the human centered *Criteria* phase before moving to the technical *Connect*, *Communicate*, and *Collect* phases. The framework also recognizes the feedback loop that refines the data collection criteria based on company's use of data in the *Consume* phase. This sequencing of phases facilitates the decision-making processes during the creation of a connected factory. *Culture* encircles all these phases and plays a critical part of the framework. As illustrated in Figure 1, the 6C Framework also functions as a process flow diagram.

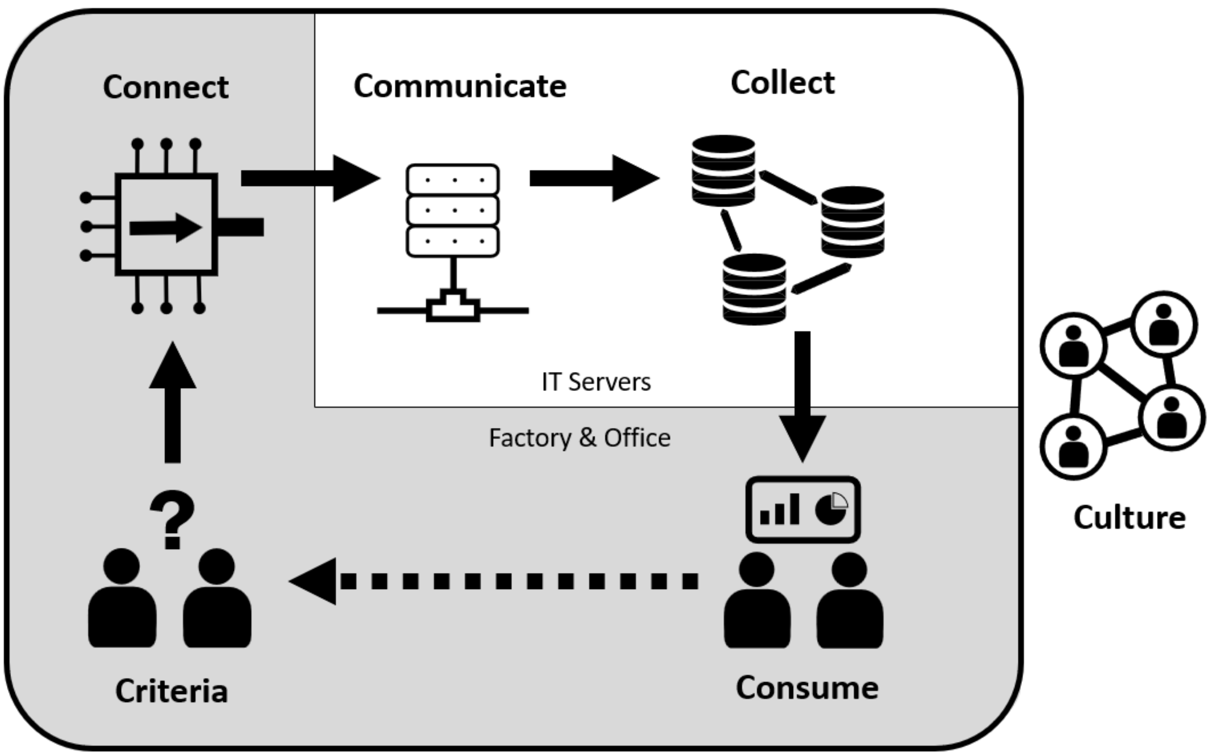

Figure 1. The 6C Framework for building a connected factory. (Artwork by author.)

FRAMEWORK HISTORY

The 6C Framework has evolved over the last decade through experience from actual I4.0 implementation in manufacturing factories. The focus during the first attempt to implement data collection in a factory was focused strictly on the technical and hardware aspects. At that time, there were only three C's in the framework: *Connect*, *Collect*, and *Consume.*[15] The primary goal then was to figure out the technical process of moving data from equipment controllers (*Connect*) to databases (*Collect*) and then ultimately the user (*Consume*). The emphasis was entirely on technical factors, neglecting the human elements. It was presumed employees would utilize data if it were accessible. This approach's flaws soon became evident.

Early attempts made it clear that data collection implementations fail without a culture dedicated to efforts. As the team solved the technical patterns of collecting and storing data, the utilization of the data lagged. Employees were not using the available data to drive decisions and communicate performance. This disconnect needed to be addressed. The data and the employees using it had to be aligned. Defining the data needs and uses had to be part of the data transformation process. Another C was added to cover this *Criteria*. Given employees played a large role in every step of the data collection process, their behaviors and patterns around the use of data greatly impact implementation success. *Culture* encircles the entire framework.[15]
The last change to the framework was made to improve the understanding and distinction within the *Collect* phase

with team members. Previously, the *Communicate* phase was grouped with *Collect*.[17] This made discussing the *Collect* phase overly complicated. *Communicate* was separated into its own phase to emphasis the importance that communication protocols and IoT Platforms (Internet of Things) have on the data collection process.

The 6C Framework was created from lessons learned during I4.0 implementations. It considers the technical and human aspects associated with change and adoption of data technologies in manufacturing. The balance of this work delves into the details of each framework phase, how the phases build on each other, and considerations that manufacturing companies can start to utilize at both strategic and tactical levels.

CRITERIA

The first C in the framework is *Criteria*. Decisions made in this phase are crucial as they drive the rest of the data collection process, from equipment and sensors used to where data is stored and how employees use the data. *Criteria* serves as the foundation upon which the entire data collection process is built. Its importance cannot be understated.

Criteria can be thought of as the design requirements of the data collection process. During this phase, all critical questions should be answered before the data collection process begins. This will avoid much wasted effort. If these questions are not addressed, the employees performing the work will have to make assumptions about what data is needed. These assumptions will never be 100% correct, which creates rework of the previously completed data collection once clarified. Common excuses heard when skipping the *Criteria* phase include this was not the right data, this was not collected at the right time, the data needed to be summarized by shift, that data should be collected at a higher frequency, this data does not show the problem, etc.

One common response from those attempting to shortcut this phase is to just collect everything. However, collecting everything is not a realistic solution. The volume of manufacturing data is overwhelming, even for basic manufacturing processes. To illustrate this, consider the simple process of sawing gating from a casting.

Gate Sawing Data

Aluminum die castings are formed by directing molten metal through gates. After the casting process, the gating system needs to be removed from the final product. In an automated diecasting cell, a large circular saw can be used to remove the gating. This section will explore the potential manufacturing data that could be generated from just this sawing process.

In any manufacturing process there are basic business metrics to be collected. These metrics include the number of parts produced or equipment uptime, which can be used to calculate parts per hour and utilization percentage. Often management needs this data enhanced with meta-data including shift, part number, machine number, tooling or fixture information, operator name, calendar day, etc. This allows data to be summarized for management to see trends by these categories. Including this meta-data is critical.

Often overlooked, the parameters set up in the equipment control are also valuable data, especially when comparing to the output results. Parameter sawing data can include the set point for revolutions per minute (RPM) and the feed rate of the shuttle. What if the typical run value for the saw is 2000 RPM, but one day it is running at 1800 RPM? This could be a trigger for maintenance on the saw. However, these results could also occur if an employee on a previous shift decided that a slower RPM would provide a longer saw blade life and made a change to the setting. This type of issue can quickly be identified by capturing the parameters setpoints from the equipment. Other manual parameter data, like the diameter or number of teeth on the saw blade, need to be captured as well. Those influence output data like amperage draw, saw blade life, and surface finish of the cut.

From a maintenance perspective, monitoring the RPM and amp draw of the motor during the cut is important. Since gate sizes vary, time-series data must be collected and analyzed. Determining the appropriate frequency for recording this data is necessary and will be discussed in the next section. Additionally, tracking manual processes occurring to the cell is essential to develop a comprehensive data story for the saw. Maintenance records need to be created and digitized to show date and time of saw blade changes, the specific saw blade style installed, and any maintenance work performed on the equipment during its life. These factors impact the RPM and amp draw during the cut. This data may be available through Enterprise Resource Planning (ERP) software or may require manual data entry. Understanding the potential data sources and being able to digitize data manually generated by employees is important.

Finally, if the goal is to collect everything then advanced sensors need to be considered. These sensors can measure the temperature of the casting as it enters the equipment and the saw blade temperature throughout the cut cycle. These temperature variables impact the saw blade's life and motor's power usage. Saw blade temperatures are also influenced by process changes such as longer time between cycles, start-of-shift warm-up, and duration of uptime on the saw. Thermal cameras can capture this type of data and provide outputs as seen in Figure 2. Infrared sensors, while more cost-effective, gather this data with less detail. The approach of collect everything lacks the necessary context and detail to determine which data collection device is appropriate for the specific problem data collection is helping address.

This gate sawing example illustrates the importance of clearly defining the *Criteria* for the data collection at the start of the process. Doing so can prevent unnecessary investment and installation of advanced sensors or time to develop connection patterns for data that may not be needed to solve the biggest problems associated with the process.

Figure 2. Example of a thermal image of saw blade temperatures. (Photo used with permission from Mercury Marine.)

Collection Frequency

Data collection frequency is an important topic in *Criteria*. Time-series data will constitute the largest volume of data generated within manufacturing processes. In discrete manufacturing, like the sawing operation, the time-series data forms a routine pattern every time a specific part is processed. This is similar to an electrocardiogram (EKG) measuring a heartbeat – a repeating pattern of data that is consistent cycle-to-cycle. This heartbeat of data exists throughout manufacturing processes. It is found in the vibration of the spindle throughout the machining operation, the fluid flow for each spray cycle of a diecasting machine, the power draw as a motor starts, and the pressure of compacting a sand mold. A key question to consider during the *Criteria* phase is: "At what frequency should this data be collected?"

Continuing with the sawing example, Figure 3 shows a process photo of the casting in a sawing operations and a modeled depiction highlighting the gate areas for each cut zone. These number zones are referenced in Figure 4, which presents actual time-series data collected from multiple sawing cycles. As the cut enters and leaves a gate area, there is variable amp draw, and power needed for the cut. This variation is evident in Figure 4. These details would be lost if a statistical value, such as an average or maximum amp draw, were calculated to represent the time-series dataset. The true picture would not be known with these representative statistics.

There is no simple answer to the question "At what frequency does this data need to be collected?" Experience from previous data collection processes can provide guidelines, but each manufacturing process is unique. Data collection frequency needs to be discussed and possibly tested during the *Criteria* phase. Poor decisions on frequency can lead to erroneous data. High frequency data collection will require special sensors, hardware, communication protocols, and data storage impacting the *Connect*, *Communicate*, *Collect* phases.

Additionally, collecting too much data can result in unnecessary data points that do not add value in analysis. To contrast, if the frequency is too low, critical maximum values can be missed. Like any engineering decision, tradeoffs occur when determining data collection processes. Collecting everything is not an option.

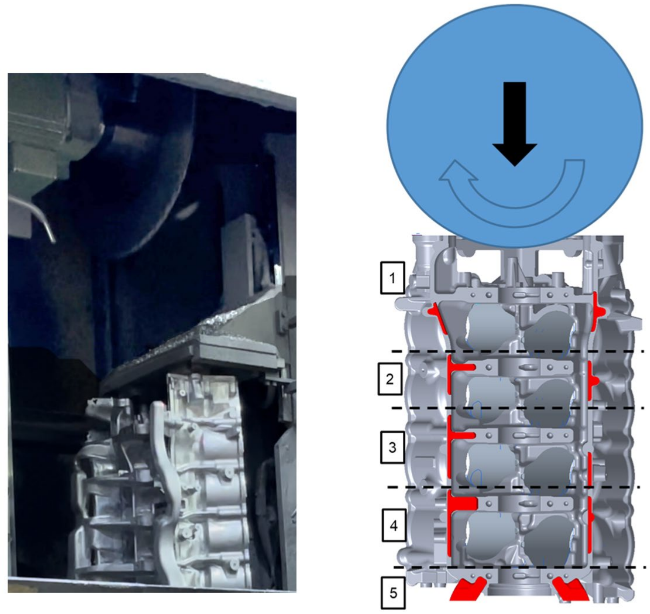

Figure 3. Gate sawing operation and simulated gate area depiction. (Photo and images used with permission from Mercury Marine.)

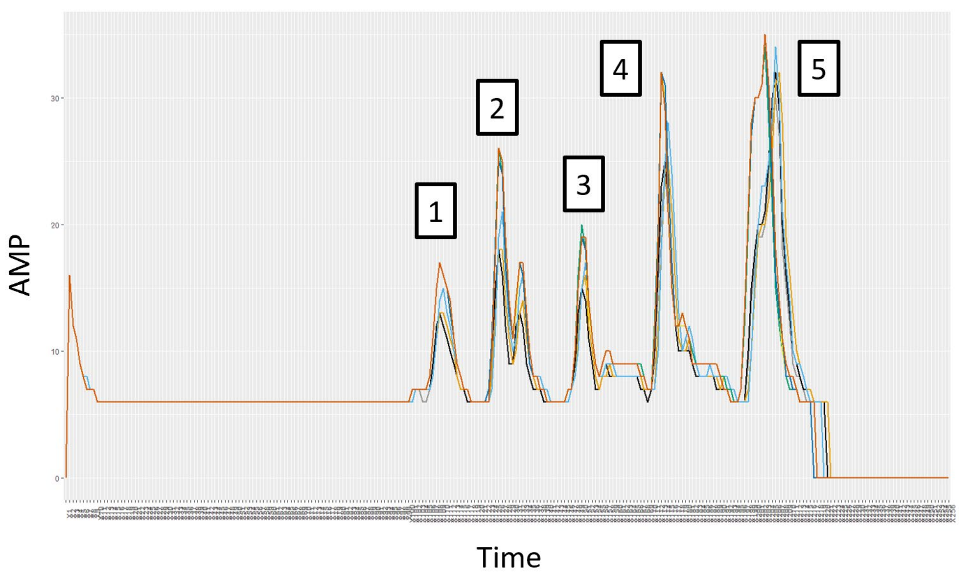

Figure 4. Amp draw data collection from sawing operations with gate areas identified.

Figure 5 illustrates the visual difference between various data collection frequencies. Data was collected at 100 millisecond (ms) data (10 Hz) during the sawing operation. This data heartbeat was processed to show how the graph would look at 500 ms (2 Hz) and 1000 ms (1 Hz) intervals. Max amp values and power, defined by the area under the curve, were calculated for each interval. The number of data points between 100 ms and 1000 ms was about 10 times smaller. With this, the max amp reading was approximately 20% less for the 1000 ms interval compared to the 100 ms, because the intervals missed the higher values compared to the higher frequency. The power values were also reduced by approximately 4% with the sampling difference. These details are presented in Table 1.

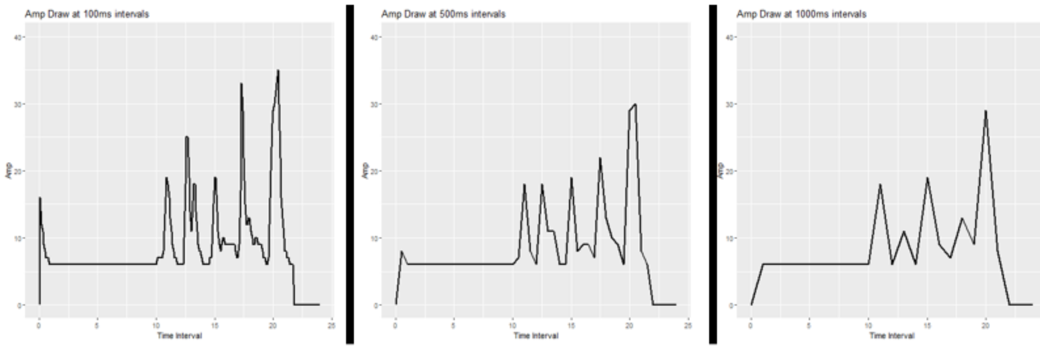

Figure 5. Amp draw data at 100, 500, and 1000 millisecond collection intervals.

Table 1. Tradeoff of Different Frequency Intervals Compared to Results

Frequency Interval	**Number of Data Points**	**Max Amp Reading**	**Power (Area under the curve)**
100 ms (10 Hz)	241	35	203.0
500 ms (2 Hz)	49	30	199.0
1000 ms (1 Hz)	25	29	195.0

Criteria Questions

This relatively simple sawing process highlights the complexity of data collection that occurs in manufacturing. Building a connected factory requires significant upfront considerations to address these types of data questions. Failure to plan during the *Criteria* phase leads to mistakes and setbacks throughout the remaining phases of the framework. Each manufacturing process and data collection is unique. Therefore, a simple framework of 5W+H (*Who*, *What*, *Where*, *Why*, *When*, and *How*) can ensure the right questions are asked for any data collection process. A cross-functional project team, including those implementing the data collection process and those utilizing the data to solve problems, should be assembled to answer the questions found in Table 2.

The *Criteria* phase is a human-centric part of the 6C Framework. While it involves many technical questions, its primary purpose it to clearly define the employees' data needs and the decisions the data will drive. This foundation ensures the data collection process is focused and efficient. Decisions are made by employees and will impact employees. Defining the details ensures the data collected will address the problem and prevent excuses for not utilizing the data.

Table 2. The 5W+H Questions for Data Collection

5W+H	Questions to Consider
Who	• Who needs to see the data? • Who will be involved in verifying the accuracy of the data? • Who will be involved in the data collection project team? • Who determines the alarm set points and triggers? • Who could manually enter data into the system? • Who is responsible for logging recipe or maintenance data/records?
What	• What type of data is collected (inputs, outputs, time-series, logic)? • What sensors and equipment need to be utilized? • What type of problems will this data solve? • What frequency does the data need to be collected? • What communication protocols are needed?
Where	• Where is the data stored? Server? Type of database? • Where will the hardware and sensors be located? • Where physically will the data be shared with the people who need it (location of dashboards)?
Why	• Why is this data important? • Why do we need to collect this information?
When	• When does the data need to be updated? • When are the trigger points used to start collecting data?
How	• How is the data delivered to users (dashboards, reports, alarms)? • How will users interact with the data collection system? • How will data be transferred? Wired? Wireless? Edge upload?

Finally, the effort and work during this phase create ownership of the collection process and the data generated. This buy-in fosters a desire among team members to see the data collection process succeed. They are engaged in the details and the future use of the data. In contrast, a project where management says collect everything and then walks out of the room is destined to fail with limited ownership. Even an incredibly talented technical team will feel lost and incorrectly assume management's objectives without the details being specified. People are a critical component of Industry 4.0, and the *Criteria* phase sets the foundation for the data collection process.

CONNECT

The *Connect* phase focuses on the technical aspects of hardware and software involved in generating data to be collected and stored. Manufacturing data sources are diverse. Sensors, humans, and software all create data. This phase includes various data patterns that can be followed based on the type of information being collected. Figure 6 illustrates different data source patterns that will be discussed.

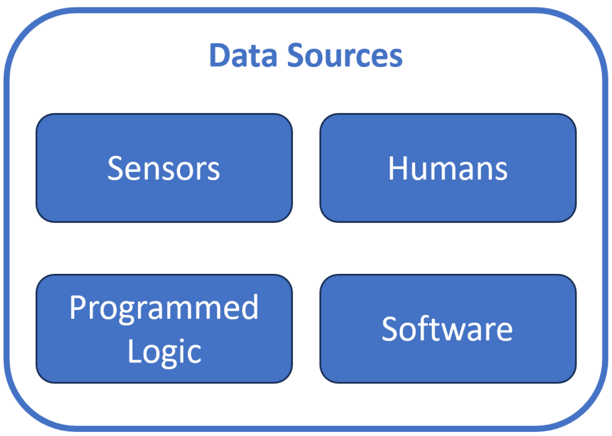

Figure 6. Data source potentials to connect.

Sensors

The most common pattern in data collections involves a sensor and a controller. This basic technology has been around for a long time. The modern thermostat, which contains a temperature sensor and a controller for a heating source, was invented in the 1800s. The technology of monitoring conditions and reacting to them grew rapidly in the early 20th century. The advent of the computer and the acceleration of the Third Industrial Revolution enabled the storage of long-term digital data from sensors. The networking of I4.0 now allows for the connection, storage, and analysis of huge volumes of data from these sensors.

A simple definition of a sensor is an input device that measures a physical quantity and provides a signal as an output. There are dozens of types of sensors, measuring inputs such as temperature, pressure, humidity, acceleration, strain, light, infrared, sound, proximity, flow rate, and color, among others. Sensors can also be built into devices such thermal cameras, which provide a two-dimensional matrix of output datapoints, or laser scans, which provide three-dimensional points of a surface of the scanned part. Data comes in many different forms, and there are many sensors to capture the values.

Functionally, the sensor detects the changing physical characteristic, and it provides the output signal to a controller. These sensors are connected to PLCs (Programmable Logic Controllers) and other types of controllers to interpret the output signal. Much of the equipment that exists today in manufacturing factories has built-in sensors and controllers to facilitate the process of the equipment. For example, in a foundry furnace a temperature sensor can detect a drop in metal temperature. The controller monitoring this value then triggers the heating element to turn on to maintain a constant metal temperature.

A significant question for companies implementing a connected factory is if these sensors and controllers can be connected to and provide data outside the equipment through a network path to a database. If the connection can occur, the project can move to the next phases. If not, hardware additions, updates, or replacements are required to gain data access. Sensors provide valuable data related to the manufacturing process. However, process data is only one part of the overall data picture within manufacturing.

Programmed Logic

The next source of data to review is *programmed logic* data. Like sensor data, this data is generated within the controller of the equipment being monitored. The key difference is that programmed logic data is not tied to a physical input that is measured by a sensor. Instead, the data is generated from logic programmed into the controller. This concept is best explained through examples.

The first example of programmed logic data is a simple part count for production reporting. Collecting manufacturing business metrics is often an initial use-case for data collection. Sometimes, a sensor counts the part as it passes through, but this is not generally the approach used. Instead, the machine is programmed to recognize a specific action in the controller indicating the process is complete and ready for the next cycle. This could be a pour complete signal coming from the foundry robot or a M30 code in a CNC program. A trigger in the system's logic informs the equipment that the part is finished and ready for the next one. The timing of this logic trigger becomes valuable information, which can be collected and converted into key business metrics to help manage the factory.

Another example is uptime data. Equipment controllers often operate in various states while processing parts. Continuing the examples, a pouring robot might be waiting to start, filling the pouring ladle, or actively pouring the metal. Similarly, a CNC machine could be waiting to load a fixture, actively machining a part, or in an idle state waiting for the next part. The controller's logic identifies the equipment's current state and tracks state transactions. This data becomes valuable when associated with time, especially the total duration between state transitions. Now, data can be converted to information telling management the total amount of time per day the machine was running compared to idle.

These are just two examples. Additional programmed logic data includes cycle time information between actions within the machine, the set points or recipe information for target values, and calculated values like heat index, which is derived from temperature, humidity, and dew point rather than a direct sensor reading. The key concept is this data is generated through the controller's logic and not a sensor value.

Humans

The third data source to connect is the employees within the company. Currently, factories have many manually tracked spreadsheets or paper charts used for various purposes, including process inspections, quality checks, visual inspections, inventory counts, and test results. For a factory implementing I4.0, this data needs to be digitized and stored into databases for long-term access and dashboarding.

In the *Connect* phase for human data, hardware and software must be considered for data entry. Often the hardware is a computer or a mobile device like a tablet or cell phone. There are numerous software options for the front-end applications for this data entry. IoT platforms, which will be discussed in *Collect*, can often provide the means to develop these types of human data entry interfaces.

These data entry screens are typically tailored to specific needs of the data being captured. Figure 7, for example, shows a screen used to collect quality checks on slurry permeability in a foundry operation. Visual inspections or quality test results are often a key type of human data. The collection pattern for human data differs from sensors or programmed logic. Selecting the right software package allows the team to easily develop applications to collect this data.

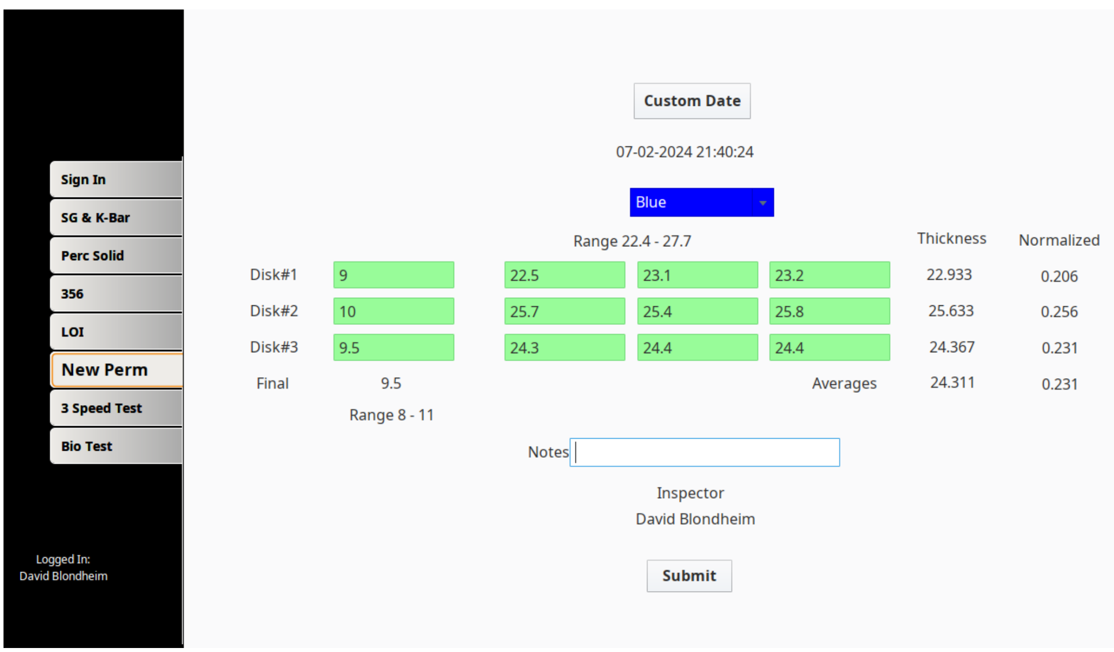

Figure 7. Employee data entry screen for quality test results. (Image used with permission from Mercury Marine.)

Software

The final data source in the *Connect* phase is data generated in different software systems. Both management and equipment software need to be considered. Management software includes systems like ERP, MES, time keeping, human resources, maintenance, sales, and scheduling, all of which are deployed by factories at different levels. Software is typically created by the OEMs (original equipment manufacturer) to control and store data from their equipment. Examples of these systems include CMM (coordinate measuring machines) software, which manages programming and data storage in one software package, and x-ray equipment that stores images during quality checks.

Additionally, with equipment manufacturers embracing I4.0 concepts and offering data services to their customers, many OEMs provide their own software that typically *Connect*, *Communicate*, and *Collect* the data from the equipment. This all-in-one solution simplifies many of the steps outlined in this framework. However, there are risks associated with this approach, such as limited access to raw data and connectivity issues with equipment brands outside of the OEM. Data is sometimes stored on OEM owned cloud servers, and the manufacturing company may only have access through pre-defined reports, without direct access to the raw data. This approach creates an income stream for the OEM, as they charge subscriptions for reporting and data storage. Years after installation, this approach can be a hurdle for factories looking to change equipment brands. Historical data could be lost or be expensive to transfer. Some equipment software is more open, allowing a company to store data on the factory's server or provide interfaces to pull data on demand without additional cost. The variety of software strategies employed by OEMs necessitates manufacturers need to learn the details of the OEM data policies and ask many questions before selecting hardware and software.

The openness of data can be overlooked when management is deciding on a software system purchase. Before I4.0, this might not have been a significant factor since systems and data were not connected. With I4.0, this becomes critical. To highlight this importance, consider the following example. Imagine ERP software used for scheduling production orders and managing inventory levels. In a connected factory, this software can interface with the equipment, allowing information such as the number of parts produced to be exchanged between systems. This can update the ERP with the number of parts remaining on the order and add new parts to inventory. The two systems seamlessly communicate this information. In contrast, a non-connected factory relies on human data entry. An operator would have to manually count the number of parts produced during the shift and enter the information into the ERP software. If the operator makes a mistake, the schedule and inventory levels will be incorrect, potentially leading to incorrect shipping dates, inventory losses, and additional production runs. Manual and repetitive tasks are prone to human error, whereas machine and computers do not make counting mistakes. The software systems companies choose are more important than ever because of this system integration.

The four data sources discussed – sensors, programmed logic, humans, and software – each have unique patterns of hardware and software to consider in the *Connect* phase. This uniqueness introduces variability that management and the technical team must understand and adapt to when implementing data collection. The scale and complexity of options for sensors, controllers, and software need to be appreciated. There are hundreds of sensor types, dozens of sensor brands, and numerous controller options, resulting in thousands of combinations of potential connection patterns. Additionally, these technologies and hardware evolve over time. A company making a sensor used today may not survive the next downturn, necessitating alternatives. Another challenge manufacturers face is that most existing factories contain equipment from various brands and processes which will require different data connection patterns. From a capital investment perspective, updating all the equipment with hardware to follow one pattern would be prohibitive. Instead, manufacturers must hire skilled technical resources capable of managing this level of variation and be fluid in implementing different patterns throughout the factory. Ultimately, the data integration should appear seamless to the end user of the data. This is achieved only through careful planning in the *Criteria* phase and understanding and execution in the *Connect* phase.

COMMUNICATE

The *Communicate* phase serves as the bridge between the hardware and sensors that generate the data and the long-term storage of the data in the *Collect* phase. Two key factors are discussed in this phase. The first is the abundance of communication protocols that exist today. The second factor is understanding *IoT Platforms*, which are software that provide the connection between the hardware and the storage.

<u>Communication Protocols</u>

A protocol, as defined by NIST (National Institute of Standards and Technology), is a set of rules such as formats and procedures to implement and control some type of association between systems, such as communication.[25] Transferring data from a sensor to a database involves using various communication protocols. The data must be packaged in a format defined by a protocol to be sent through the system. There are protocols for addressing where the data is being sent within the network. Security, routing, and queuing protocols ensure safe and proper delivery of the data on this network. Transportation protocols define the flow of data through wires or airwaves. While all these protocols are critical to the success of data collection implementation, the focus in this work will be on two specific types of protocols.

The first protocol is the logistical aspects of moving data from the source to the storage. A decision must be made between wired or wireless connections to the hardware gathering the data. Wired connections offer robust solutions with better speed, lower latency, higher bandwidth, and more control and security over wireless connections. However, each new equipment requires a physical connection, taking time and cost to run network cables. Wired connections also face challenges with cable clutter and limited mobility. On the other hand, wireless connections eliminate the clutter and connection time but present challenges with security, bandwidth, speed, and throughput depending on the application. Various protocols, such as Wi-Fi, cellular, and Bluetooth, can be used.

Manufacturing companies will eventually adopt a hybrid approach, utilizing both wire and wireless solutions. For critical, high-speed data collection, wired connections are preferred. For low speed, non-critical data, such as human-entered test results, wireless communication via a laptop will suffice. Each type has its trade-offs. This wired versus wireless decision should be made during the *Criteria* phase, ensuring the team can effectively execute the plan in the *Communicate* phase.

The second protocol manufacturers must consider is the application protocol, which dictates how devices communicate with each other. In the industrial data collection space, many commonly used protocols exist. A few different protocols include OPC UA, MQTT, REST API, HTTP, MTConnect, Modbus, BACnet, and Profibus. Each protocol was developed to help address specific shortcomings of the others. As with all engineering problems, trade-offs occur between choices. For example, low file size and bandwidth is extremely important in cellular data connections, while speed and robustness are more important in other scenarios.

For most manufacturers, the various equipment installed in their facilities will use different protocols. BACnet protocols are heavily used in building automation and heating, ventilation and cooling (HVAC) systems. Be prepared to use these protocols when trying to connect to HVAC systems to collect real-time energy use or temperature data within the factory. MTConnect is a popular protocol used in CNC machining equipment. OPC UA or Modbus are commonly used by industrial PLCs found in manufacturing equipment. MQTT is a lightweight protocol often leveraged for wireless data collection applications.

There is a large variety of communication protocols used in manufacturing today. Hiring skilled technical resources is a requirement to manage the variation of protocols. These employees benefit from the slow evolution of the different communication protocols. Additionally, IoT platforms, which will be discussed next, often contain extensive driver libraries, facilitating easy connections to these protocols. The challenge, and where the talented technical team helps, is to ensure the data is correctly set up on the different hardware devices to enable successful communication between the IoT platform and the hardware.

IoT Platforms

An IoT platform is a software package that connects devices to networks and manages the data that is generated. IoT platforms are the bridge between the hardware and the storage of the data. It provides a connection between OT (Operational Technology) and IT (Information Technology).

The IoT platform also provides control and logic to the data collection process. This is best explained with an example. Imagine a foundry with a furnace holding aluminum at a liquid temperature. Within this equipment, a temperature sensor in the metal is connected to a PLC. A tag, or variable, represents the temperature value stored in the PLC. The PLC can read this sensor near real-time and adjust the furnace controls, turning the heating element on or off to maintain the temperature. The PLC is connected to the IoT platform through the network. The IoT platform communicates with the PLC and has access to all the PLC tags. The IoT platform can see the sensor values changing in the PLC. Additionally, the IoT platform has access to the database server that stores all the historical process data.

With this bridge in place, the IoT platform must be configured to trigger data transfer based on the established criteria. If the requirement is to collect data every five minutes regardless of the process, the IoT platform would populate the database with rules created to write the tag to the database with a five-minute timer trigger. Alternatively, if the criteria specify collecting temperature data right before the next robot takes a ladle of metal out, a different rule would be created in the IoT platform. This rule would monitor the metal temperature tag but only trigger a data write when a separate tag controlled by the robot is activated. Instead of having data points every five minutes and estimating which temperature is closest to the casting process, a set trigger from the casting process controls the data generation through the IoT platform. The *Criteria* should define these triggers and data collection rules to ensure the IoT platform generates the desired data. This example illustrates the important logic and control that IoT platforms provide in the data collection process.

Finally, selecting an IoT platform is probably the most significant decision management must make when starting to implement data collection at a factory. As with other aspects discussed, many different software packages can serve as an IoT platform. Each has advantages and disadvantages. Unlike sensors, controllers, and communication protocols, once an IoT platform is chosen, a factory typically only uses that one platform. It becomes a long-term commitment. Employees need to be trained, and many rules and triggers will be programmed into the IoT platform that will be difficult or impossible to transfer to other software in the future. This makes the

IoT platform a critical management decision for the data collection process.

In the *Communicate* phase, the device has been bridged and can transfer data to storage. The OT and IT systems are connected. All the effort to gather this data is about to be stored in a location discussed in the next phase.

COLLECT

The *Collect* phase focuses on storing data within IT systems. This is a culmination of all the effort of the previous phases of framework. The company plans the data collection process and details (*Criteria*), generates the data with sensors or hardware (*Connect*), and utilizes protocols to transfer the data (*Communicate*), with the goal of storing the data for immediate or future use (*Collect*).

The tools used in the *Connect* phase are typical skill sets of IT professionals. Data is stored in databases, which are managed on servers. Databases offer a structured method for storing data. Database tables are created to hold specific types of data. These data types include numbers, dates, time, characters, and others. The IoT platform organizes the data communicated from the hardware in a structured manner to be written into these database tables. This is referred to as structured data. In contrast, unstructured data is used to store data that does not fit into structured databases format, such as text files, video, and images. Both storage options will be discussed within this section.

Structured Data

The data generated by sensors often ends up in a structured format. As companies embark on their data collection journey within I4.0, they will need to understand structured or relationship databases.

A database is usually stored on a server dedicated to managing databases. A server can hold multiple databases. A database is a collection of tables, which contain columns and rows of information. Each row represents a record with a unique identifier and all associated data. Each column contains individual data attributes of that record.

Structured databases are called relationship databases since they establish relationships between different tables using primary and foreign keys. To illustrate this, consider two tables used for sales order software. One table, the *CustomerTable*, contains all the information regarding the customer, such as their name, shipping address, primary contact, phone number, payment terms, and so on. This table includes a primary key, like a unique number or customer ID that increments with each row. The second table, the *OrderTable*, contains information about each order placed with the factory. During the order process, the customer is selected in the software. Instead of the *OrderTable* duplicating all the customer information (name, phone number, address, etcetera), it includes a single column for the *CustomerID*, which is a foreign key that links back to the unique customer identifier in the *CustomerTable*. Relationship databases can join this information together. Figure 8 provides a visualization of this concept.

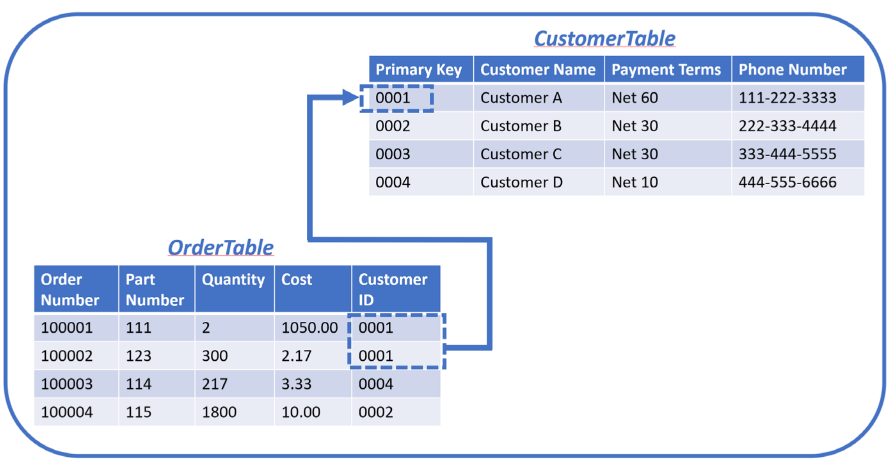

Figure 8. Example of a relationship between tables.

When collecting manufacturing data, there are two primary data storage philosophies to consider: *wide tables* and *tall tables*. A wide table has records that span multiple columns with related information. For example, a quality inspection entry might record the test ID number, date, test type, quantity checked, part number, and results. Wide tables result in fewer rows of data but more columns. In contrast, a tall table contains one of these observations per row, linked by the test identification number. This means fewer columns of data but many more rows. Figure 9 provides a visual comparison of the structure differences between tall and wide tables for two example inspections.

Tall and wide tables contain the same information but structure it differently. There are positives and negatives for each structure. People are familiar with wide table structures, as data is often presented this way in daily work. Wide tables can be challenging when changes to the schema occur. Imagine if a new observation for the inspector's name needs to be added to the tables. In a wide table, this will create a new column, and all previous rows of data stored into the database will have empty or NULL value associated with it, since this data was not known at the time of entry. In a tall table, new variables can be added or removed at any time without changing the data model of the table. It is just another row of data stored with different information. Another benefit of tall tables is that software used to create charts and plots of large datasets, which will be discussed in the *Consume* phase, may prefer, or even require, data in a tall table format.

Unstructured Data

Text files, PDFs, emails, images, videos, and audio files are all examples of unstructured data. Typically, factories begin with structured data storage but eventually need to store advanced datasets provided by unstructured data. These datasets often offer valuable insights for advanced analytics, including AI tools like large language models (LLM), making them important to consider with I4.0 implementation.

Unstructured data storage is also referred to as *blob storage* or sometimes a *data lake*. These are specialized storage devices capable of handling file types like storing images, videos, and more. Blob storage and data lake are not fully interchangeable terms since a data lake can contain structured and unstructured databases. Unfortunately, unstructured storage is often more difficult to search, sort, and analyze when compared to schemas of relationship databases.

Wide Table

Test ID	Date	Test Type	Quantity Checked	Part Number	Results
T01	2/2/2024	Hardness	1	123	Passed
T02	2/17/2024	X-Ray	10	234	Failed

Tall Table

Test ID	Variable	Value
T01	Date	2/2/2024
T01	Test Type	Hardness
T01	Quantity Checked	1
T01	Part Number	123
T01	Results	Passed
T02	Date	2/17/2024
T02	Test Type	X-Ray
T02	Quantity Checked	10
T02	Part Number	234
T02	Results	Failed

Figure 9. Wide and tall table examples.

It is also important to consider semi-structured data that can exist in unstructured storage. Semi-structured data uses tags and metadata to help organize information within the file. Examples of semi-structured file types include XML (Extensible Markup Language) documents and JSON (JavaScript Object Notation) files. An example of a JSON file for test T01 found in the tall table from Figure 9 can be seen in Figure 10. The structured nature of the data can be seen without the need for a formal database.

Structured, unstructured, and semi-structured provide options for storing different types of data. As a manufacturing company's data collection implementation matures, it will gain experience with all these storage types. Initially, the focus will be on structured databases, which can handle all but the most complex data. This initial focus is important because many commercially available software solutions for reporting and dashboarding are designed to work with relationship databases. These details will be reviewed next in *Consume*.

JSON File Example


```
[
 {
  "Test ID": "T01",
  "Variable": "Date",
  "Value": "2/2/2024"
 },
 {
  "Test ID": "T01",
  "Variable": "Test Type",
  "Value": "Hardness"
 },
 {
  "Test ID": "T01",
  "Variable": "Quantity Checked",
  "Value": 1
 },
 {
  "Test ID": "T01",
  "Variable": "Part Number",
  "Value": 123
 },
 {
  "Test ID": "T01",
  "Variable": "Results",
  "Value": "Passed"
 }
]
```


Figure 10. JSON file format example.

CONSUME

Collecting data without utilizing or learning from it is known as the collector's fallacy. The *Consume* phase of the framework aims to prevent this fallacy by transforming gathered data into actionable information and knowledge. Employees then use this information to solve problems and improve the business.
The *Consume* phase involves technical aspects related to the programming of the software used to interface with the databases and manipulate the data, but this is not the most important part of the phase. The core function of this phase is ensuring that people are utilizing the data generated to make decisions and act on improvements. People are central to this phase of the framework.

A well-defined *Criteria* will make the *Consume* phase straightforward. The exact type of data should be specified, with plans clearly outlined for dashboards, reports, and alerts. Additionally, employees should be highly interested in obtaining this information to drive decisions. Defined *Criteria* will ensure the immediate utilization of data in the *Consume* phase, avoiding any excuses for not using data. The transfer of this data to information can occur through real-time dashboards (RTD), historical reporting, alarming and alerting, and predictive analytics.

Real-Time Dashboards

A real-time dashboard (RTD) visually displays key metrics and the status of equipment and processes within the factory. A RTD automatically updates with the latest information, enabling operators, supervisors, and management to understand the factory's real-time performance. Although RTDs can frequently refresh the data, the speed of the updates varies. Some RTDs are updated every second, while others, such as those showing product volumes with long process cycles, may update every few minutes or even on the hour. These dashboards can show more than current values. They often contain recent historical data where trends or patterns can aid in the decision-making process.

There are numerous software providers that offer real-time dashboard solutions. Often, IoT platforms include a module for creating RTDs. Access to tag values in PLCs makes it straightforward to provide this data to visuals used in dashboards. Management must invest time in evaluating RTD software options, as they are a critical part of the data journey. These dashboards are usually displayed directly on the factory floor. They are the most visual representation of the data collection process to employees. While RTDs are often one-directional, they do not have to be. Employees can monitor data but also have the capability of entering notes or comments into these dashboards, turned data entry screens. Design and implementation of these dashboards is critical for success. An example of an RTD can be seen in Figure 11.

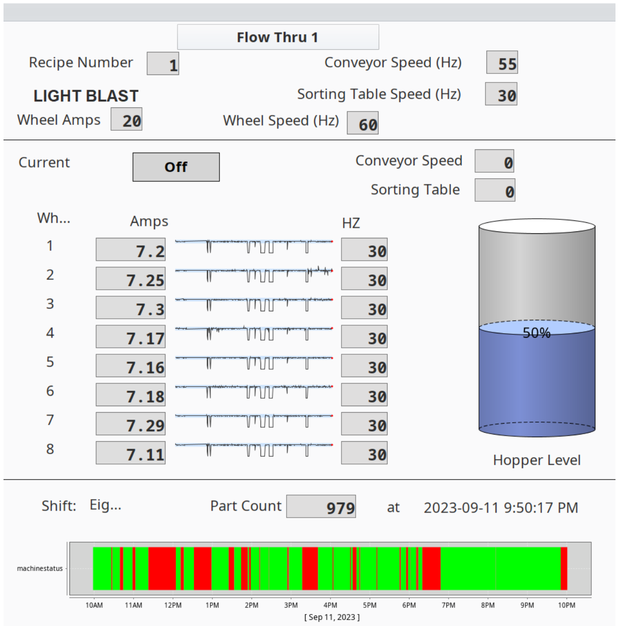

Figure 11. Example of a real-time dashboard. (Image used with permission from Mercury Marine.)

Historical Reporting

Real-time dashboards are important for the immediate, tactical decision-making that occurs daily. In contrast, historical reporting provides insights into long-term trends. Historical reporting shows performance through time, such as the daily max amp draw on a HVAC motor for the past six months or business metrics like monthly scrap rates or uptime for the past five years. The objective of historical reporting is to understand past performance and identify trends. It helps to validate improvements made or pinpoint areas needing further attention. Historical reporting is important for monitoring the strategic health of a company.

Manufacturers have numerous software options for historical reporting, each offering various methods for delivering reports. Many systems provide web-based reporting system similar to RTD, where users can log in and view different data visuals. Another common option is scheduled email reports, which are sent out to employees at certain dates or times. The report will appear as an email or attachment. An example of a historical report is seen in Figure 12.

One of the strong links with the *Collect* and *Consume* phases is highlighted through historical reporting. This data is often stored in a relationship database. Part of the software process involves requesting data from these databases for the charts, reports, and other visuals. This request process is called a *query*.

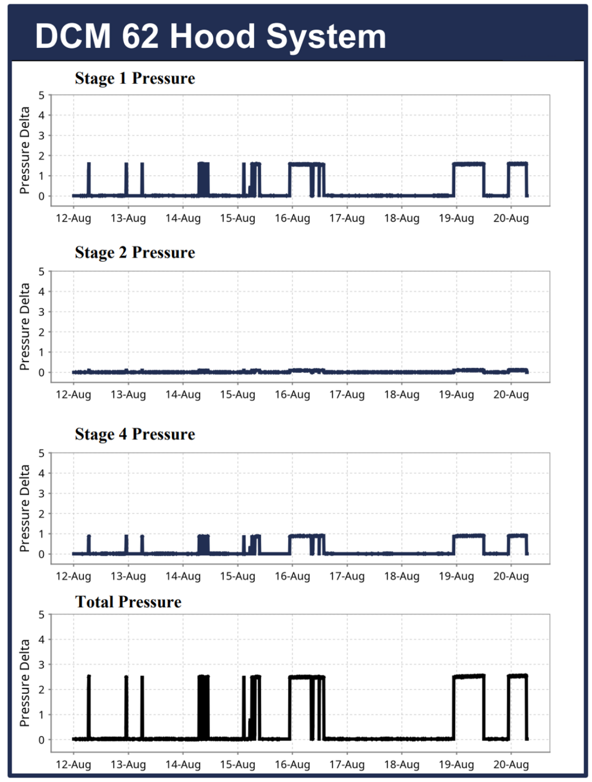

Figure 12. Historical report example. (Image used with permission from Mercury Marine.)

There are different languages used to send queries based on the type of relationship database. SQL (structured query language) is one such programming language used for processing the data within a relationship database. Understanding SQL and other database commands is essential for unlocking data from databases and making it available to be used.

Alarming and Alerting

Alarms and alerts are indications of an abnormal condition in the process or equipment. For decades, alarms have existed in manufacturing equipment. However, these alarms were limited to those operators in sight or earshot of the equipment. When something went wrong, the operator would need to recognize and response to the alarm, typically calling supervisors or maintenance team members to assist and troubleshoot. Manual reactions and communication are the norm in unconnected factories.

With data connections of I4.0, companies can leverage automated data analysis to monitor and generate alarms and alerts throughout the organization. Companies are no longer reliant on employees visually monitoring and responding to data at equipment. Instead, a computer program evaluates data as it comes into the system, comparing the new data to predefined rules or thresholds.

Once the rule is broken, like the temperature of the process gets too high or a holding tank is close to overflowing, the computer triggers a notification sent to the relevant personnel across the organization. This automatic analysis becomes critical as data connections grow to cover the entire factory with all of its equipment.

It is important to understand the difference between alarm and alert, as these terms are often incorrectly used. The International Society of Automation (ISA) publishes the standard on alarming in ANSI/ISA-18.2 *Management of Alarm Systems for the Process Industries*. According to this standard, an alarm requires an immediate response, while an alert is a notification that may not require a timely response.[26] The key difference is whether immediate action is needed by the employees.

Alarms and alerts can be visible or audible notifications of process or equipment anomalies. Blinking lights or buzzers can be used for alarms or alerts in critical areas in the factory. Sometimes, the people requiring notification may not be physically near these areas. This is where I4.0 technology can be utilized. Dashboards in a maintenance area could start flashing when any motor within the factory starts to fail. Similarly, a dashboard in the manufacturing engineering office could highlight failed quality checks that require an engineer to verify a process. The technology exists to push alarms and alerts via emails, texts, and even phone calls.

The first step of sending alarms and alerts is developing the rules for the software to follow. IoT platforms often include modules for this purpose, but other software options are also available. Generally, the software needs access to the tag or sensor directly or via the database where the value is stored. It monitors the value and compares it to the predefined rules. These rule usually include threshold values that trigger notifications when exceeded. Alarming and alerting software often have additional functionalities, such as tracking response times, escalating alarms if there is no response, and maintaining a history of alarms. This data can be used to improve response times, thereby reducing downtime or process inefficiencies within the factory.

The benefit of these notifications is the employees only consume the data when there is an action needed. One challenge with I4.0 data is the amount of information that can be automatically distributed. People become desensitized and unresponsive to these automatic notifications if there are too many of them. The company improves its chance of success by developing a robust alarming policy that focuses on true anomalies that require a timely response. Employees will understand the alarms are critical to the business, with meaningful thresholds to drive immediate action. This approach avoids nuisance alarms based on poorly designed thresholds and over-notification. Additionally, the software will provide metrics on responses and give management the opportunity to measure and drive improvement.

This is a basic introduction on the topic of alarming and alerting, a critical method of utilizing data in the *Consume* phase. Additional effort and research are needed to be fully knowledgeable regarding alarms and alerts. A robust understanding of the equipment and values that drive process anomalies are also required. Re-engaging employees after they experience a poorly designed alarming and alerting system is difficult. Therefore, this method of consuming manufacturing data needs to be executed well from the start.

Predictive Analytics

Artificial intelligence (AI) and machine learning represent the future and pinnacle of data consumption for I4.0. Predictive analytics utilize historical manufacturing data to identify patterns, enabling future predictions or the detection of process anomalies. AI and ML create these prediction models without being explicitly programmed.

While entire books cover topics of applications of AI and ML, this section highlights the significance and types of predictive analytics. Generative and applied AI will be defined and examples discussed to provide a general overview on the topic of advanced analytics. AI and ML are poised to revolutionize the future of manufacturing, but these tools are complex and require substantial effort and data to yield results.

Generative and applied AI are important subsets of the broader field of artificial intelligence, each serving different purposes and having distinct functionalities. Generative AI creates new content, typically text and images, based on learning patterns from existing datasets. In contrast, applied AI focuses on solving specific, practical problems by applying machine learning algorithms to real-world data and tasks. Both cases require datasets to be utilized to help train the AI, though the type of data and problems they address differ significantly.

Large language models (LLMs), a tremendous growth area of generative AI, are used to understand, summarize, and generate new content from datasets. As discussed in the *Collect* phase, unstructured data – such as PDFs of equipment manuals, emails of daily production notes, or previous work instructions – can be fed into LLMs to generate summaries and insights into the factory. A typical interface for LLMs is a chatbot. Using chatbots with LLM can eliminate the time-consuming process of humans sifting through all the unstructured data generated in a factory.

Applied AI focuses on identifying patterns in data to make predictions about specific processes it was trained on. Classic examples of applied AI include utilizing equipment data to predict machine failure or part quality or applying computer vision to images. Process data is collected and used to train these algorithms, which then make future predictions based on current process conditions. Ideally, applied AI enables predictive maintenance to be performed before a motor fails, optimizes a process to avoid scrap, or helps complete visual inspections to with cameras to improve customer quality. Applied AI uses various types of machine learning algorithms, ranging from easy-to-understand models to complex 'black-box' algorithms, where the reasoning behind prediction is not transparent.

Based on the author's experience, real-time dashboards, historical reports, and alarming/alerting should be prioritized over predictive analytics for manufacturing companies in the *Consume* phase. The data that feeds into predictive analytics must be validated and well understood. The adage "garbage in equals garbage out" is especially true for predictive analytic tools. Real-time dashboards and historical reporting are the first steps in validating manufacturing data. Additionally, these tools enable management to drive company improvements that are easy to explain to employees. The reasoning behind some AI and ML models is not always clear and can be difficult or impossible to explain. While real-time dashboards and reports are easily scalable across diverse equipment, predictive analytics may require specialized data, programming, and work on each piece of equipment separately.

This lower prioritization does not imply that predictive analytics are unimportant – quite the opposite. Predictive analytics is likely the most important use of data consumption in the long term for factories. However, implementation involving the data, software, algorithms, and prediction can be complex and challenging. Early and quick wins are crucial when starting a data collection implementation. Predictive analytics is unlikely to provide quick wins and may become a resource drain on the technical team with little useful results to show initially.

Additionally, as will be discussed in the next phase, employee acceptance must be considered. Employees who struggle to utilize foundational dashboards and reports to improve the organization will not blindly embrace AI and machine learning. There will be challenges in using the data that predictive analytics provides. *Culture* is important in the overall success of not only the *Consume* phase but the entire framework.

CULTURE

The technologies of Industry 4.0 and data collection are highly specialized. Sensors, controllers, software, and coding represent technical decisions and challenges for a company. Research publications and studies often focus on these technical components.[27] However, technology is only half the conversation. The employees are the other half. From experience, people can sometimes be the more challenging piece of the puzzle.

Historically, since the late 1980s to the start of the 2010s, the United States has experienced a consistent increase in labor productivity in the manufacturing sector because of the technologies brought about by the third industrial revolution. The US Bureau of Labor Statistics publishes the labor productivity metric. Interestingly, since 2010, productivity has plateaued. This is illustrated in Figure 13.[28] As mentioned in the introduction, Industry 4.0 was coined in 2011. This timing is ironically near the start of the labor productivity index plateau. This initially seems counterintuitive, given that industrial revolutions typically bring about significant changes and improvement to manufacturing. Technology may be available, but productivity gains only come when people are willing to adopt and utilize the technology.

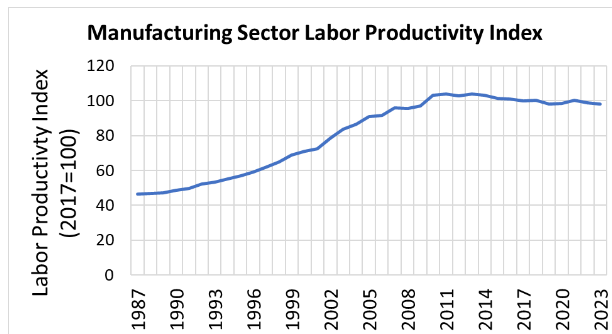

Figure 13. U.S. Bureau of Labor Statistics-Manufacturing Productivity Index. [28]

This conflict with technology adoption is seen throughout history. In the early 1800s, the Luddites were textile weavers who destroyed the mechanized looms, fearing the technology would cost them their jobs.[29] In 1831, 150 to 200 tailors rioted and destroyed dozens of sewing machines believing the machines would lower their wages and increase unemployment.[30] The earlier story on Henry Ford showed the workforce leaving assembly lines in masses due to changes in the working conditions. Ford was only successful at changing this culture after a dedicated effort was made at changing pay and employee hours. This fear of technology is not just a thing of the past. Terms like "computerphobia" or "cyberphobic" were used to describe people with significant fears of using or touching computers. A study from the early 1980s found that about one-third of 500 college students and corporate managers had some level of cyberphobia.[31] Fast forward to today, and it is hard to open a newsfeed or listen to the news and not hear about the fear of AI ending humanity.

People do not change easily; it takes education, time, and effort. Manufacturing companies should expect these same needs when adopting I4.0 and data collection. Companies will only achieve the benefits of streamlined decisions-making, improved productivity, and increased quality when their employees fully embrace the use of data. People have been integrated into many of the different C's of this framework already. *Criteria* and *Consume* are phases dependent on people in the process. The last C, *Culture*, examines the overall human aspect of implementing I4.0 within a company. A company's culture can be defined as the human behaviors, norms, and patterns used within a given company.[27,32] Within the company's culture is a subset of cultural behaviors associated with the use of data and data-driven decision making.

Culture is a difficult topic, which is why it is not always covered in publications on the technical implementation of I4.0.[27] There is no single correct method for implementing I4.0. What works for one company may not work for the next. Manufacturers like to follow steps to produce a known outcome, like a standard operating procedure to achieve the desired product. This does not exist with culture, and, therefore, it does not exist when implementing I4.0 technologies. There is no set of instructions to follow and arrive at the same destination as a competitor in the industry. This is why many I4.0 implementations have not been successful over the past decade. This framework is needed to guide management down the right path, allowing them to adopt the implementation right for the culture of their company.

Even without a set of detailed directions, there are concepts that companies need to consider to ensure their compass is pointing in the right direction for their data journey. The first of these is recognizing and addressing the skill set gap that likely exists in manufacturing companies. Throughout this framework, there have been numerous mentions of sensors, hardware, communication protocols, databases, reporting, coding, and software. These terms involve OT and IT skill sets that are not commonplace in manufacturing companies, especially small- and medium-sized manufacturers. In these companies, people often serve multiple roles. For example, the manufacturing engineering manager might also help run the IT department for a small company because he or she has an interest in computers. There is a chance that the manager may be the entire IT department. Similarly, a recent graduate might oversee a small company's IT system because they have experience setting up servers for gaming in college. The resourcefulness of manufacturing companies is always amazing. However, the tradeoff is these employees lack the specialized skill sets and time to fully develop and advance the IT and OT components necessary for I4.0.

One of the first realizations experienced when implementing I4.0 is the need to understand programming languages and databases. Data flows into and out of databases with software. No coding means no data. One way to improve this skill set on the I4.0 team is to hire college graduates with computer science degrees. This approach leveraged their education and skill sets to develop and program the necessary software to extract the data from the manufacturing equipment. They are trained in manufacturing, and in turn, the manufacturing team is trained in IT. Both educations were essential for success. Hiring skill sets outside of the norm can be challenging for small- and medium-sized manufacturers due to budget constraints. However, management must be resourceful to overcome this challenge to aid in the rapid adoption of I4.0 technology.

Another concept that will help companies improve their data culture is ensuring management leads with data-driven decision making rather than gut feelings. Data needs to become part of the daily conversation at all levels of the organization. Company-wide metrics need to be defined and turned into department metrics, which are cascaded down to operating facing cell-level metrics. These cell-level metrics need to be tactical and easily understood by the operators. Simplified metrics, such as the number of parts produced or total uptime hours, are preferred over the metrics managers typically utilize, like efficiency or uptime percentages. The goal is to provide a tactical measurement that operators can relate to and influence directly. Increasing efficiency by 5% can be confusing or difficult to understand for some people. However, most people can understand the value of producing two more parts a shift can create. These two metrics may be identical, but extra parts can be physically handled by the operator, while a percentage is a calculation held in the mind. The physical nature of manufacturing creates the need for tangible metrics for operators.

Having the goal-setting system cascade and disseminate targets is the first half of the problem. The second half involves building a culture of responsibility and accountability for achieving these targets. Regular reviews and communication of metric results are needed at all levels of the company. This creates a symbiotic relationship within the organization. The front-line employees must be held accountable for delivering on the metrics set by their supervisor. In turn, the supervisors are held accountable by the managers, managers by vice-presidents, and so on. Responsibility though the opposite direction: vice-presidents must ensure that reviews are conducted with managers and that they have the necessary resources to succeed. Managers are responsible to supervisors to remove barriers. Supervisors are responsible for aiding front-line employees. These are two-way streets that need to foster dialog and conversation around the metrics.

This drives an understanding of what the metrics aim to achieve and, more importantly, integrates these data-driven metrics into the company's norms and behaviors, becoming part of the company's culture. This does not happen by accident; it takes consistent effort from all levels of the organization.

Another critical concept to consider is the significant change in daily work that some employees will experience when using data. As a trained engineer, using data to drive decisions was always a norm for the author, however, this is not the case for all people in manufacturing. For many factories, data has never existed or been available to solve the daily problems. Historically, supervisors and managers made decisions with limited or possibly no data. These decisions built experience and confidence in their decision-making ability through the years, and possibly decades, they worked. This approach is disrupted within a data-driven I4.0 company. The data now exists and is readily available to assist with many manufacturing decisions like scheduling operators based on efficiencies, showing the performance of a dozen machines real-time on the supervisor's mobile phone, identifying quality issues as the occur, and communicating performance to all levels of the organization in real time. The role of many supervisors and managers will drastically change with the increased use of data. While it is unlikely the Luddites will return with pitchforks to destroy the I4.0 data collection systems and servers, the amount of change in work content for employees, supervisors, and managers will be monumental. As history has shown, not all people adapt well to such changes.

The interconnectedness of data and breaking the data silo mentality is another necessary cultural shift. The third industrial revolution began generating data, but the networking and connected systems did not exist yet. Data silos were created because they were the only option. ERP software and its data did not communicate with the timekeeping software employees punched in and out daily. The motor performance data from the building management system did not share information with the maintenance software. The dimensional data from the quality CMMs was isolated from everything else. As networks were built, the ability to connect and communicate this data became available. Today, many companies still behave as if the data from these systems is siloed and should not be connected. The policies of some software providers can hamper this connectivity.

By leveraging the advantages and desires of company to implement I4.0, software providers are building out modules and storage options they can sell as a service with monthly costs charged. Some companies are forthright with the data, providing it to the customer to leverage. Others, however, are less scrupulous and try to maximize sales of software while holding the data hostage. Management of manufacturing factories needs to have a policy and approach to selecting and integrating the multitude of software systems that help run a factory. Will management accept a timekeeping software vendor that only allows data to be accessed through its reporting? This could prevent the connection to an ERP system that could automate inventory allocation, avoiding a manual process for employees. What if this software is already in place? What will be the management's behavior and pattern in these cases? Management is responsible for defining the culture of sharing and connecting data. The approach with vendors also cascades to connecting and sharing data between departments within the company. Transitioning away from a siloed mentality of data is an important part of the culture of data.

To support the data journey, training needs to occur within the company. This includes technical training on new software packages and the hardware that will be utilized. Additionally, training must cover general data and quality concepts that may not be currently used. With real-time streaming process data, quality concepts like statistical process control charts (SPC) and capability studies can be performed continuously on processes with every new piece of data captured. Advanced analytics skills, such as machine learning, should be reviewed at all levels of the organization, so employees understand why these tools are being used, even if they are not programming the algorithms themselves. In some cases, basic mathematical concepts like calculating percentages and understanding graphs may need to be part of a training program for employees. The need and level of training will differ between companies. Recognizing culture is driven by behaviors, and training is an important method to reshape employee's behaviors and willingness to change.

Culture drives the adoption of technology. Changing culture requires management to provide consistent effort through time. Change will not be immediate. The questions and comments in this section, like all sections of the framework, are designed to make the management think deeply about the steps needed for the adoption of I4.0. The examples shared are a subset from the author's experiences. There are many other potential strategies to influence behavior and change culture that can be researched and implemented. The objective was to introduce the topic and emphasize the significant influence of *Culture* on the adoption of I4.0, though each culture is unique.

CONCLUSIONS

The 6C Framework offers a structured approach to translate the concepts of I4.0's factory digitization into practical applications. The phases – *Criteria*, *Connect*, *Communicate*, *Collect*, and *Consume* – are essential for planning, executing, and benefiting from data utilization in manufacturing. This process is influenced by the data *Culture* of the company, encircling all the phases in the framework.

The framework begins with people coming together to plan the data process. *Criteria* defines the need, giving clarity to the needs in the other phases. This includes identifying the problems the data will help solve once the implementation is complete. With this foundation in place, the data collection moves into the hardware and data definition phase, known as *Connect*. The goal of this phase is to execute the hardware or software creating the data. Once the data is available, *Communicate* outlines how it moves through the network via communication protocols and an IoT platform. The IoT platform delivers the data to long-term storage in *Collect*. The type of data and how it is structured are important items to address in this phase. With the data in place, the company can leverage it to make improvements in the *Consume* phase. Real-time dashboards, historical reporting, alarming and alerting, and predictive analytics are tools that can be used to deliver the data to the employees. *Culture*, or how the employees interact and behave with data, influences all these steps.

Management can use this framework to guide the decisions necessary for implementing a large-scale data collection system within any size factory. The 6C Framework is more than just a strategic tool. The same phases can be tactically applied to any individual data collection problem. The goal with this framework is to translate the theory and concepts of data collection into tangible, actionable phases that any company can implement.

Building a connected factory comes with challenges. Different processes, equipment vendors, age of equipment, and software require various technical implementation patterns to be developed to harness the manufacturing data. Additionally, employees need to be actively engaged in planning the criteria and consuming the collected data. The human aspects can easily be ignored given all the technical challenges and decisions needed when starting I4.0 implementation. For the engineers and technicians implementing this I4.0 data collection, the focus on technical decisions and equipment testing can overshadow the people component. This oversight is a recipe for failure, which has plagued many implementation attempts in the industry thus far.

Successful data implementation involves integrating the technology (OT), the data (IT), and the people. The purpose of data in Industry 4.0 is to optimize decision-making and enhance the productivity of employees involved in manufacturing. Data serves as a starting point in a long-term Industry 4.0 improvement journey. When leveraged properly, data can revolutionize manufacturing. This 6C Framework provides a guide for manufacturers to embark on the digital transformation journey.

REFERENCES

1. Hoyer, C., Gunawan, I., Reaiche, C.H., "The Implementation of Industry 4.0 - A Systematic Literature Review of the Key Factors," *Syst Res Behav Sci.* (2020) Jul;37(4):557–78. https://onlinelibrary.wiley.com/doi/10.1002/sres.2701 (Link last accessed 01-16-25.)
2. Pfeiffer, S. , "The Vision of "Industrie 4.0" in the Making—a Case of Future Told, Tamed, and Traded," *Nanoethics* (2017-Apr);11(1):107–21. http://link.springer.com/10.1007/s11569-016-0280-3 (Link last accessed 01-16-25.)
3. "Industry 4.0 after the initial hype. Where manufacturers are finding value and how they can best capture it." McKinsey Digital (2016) https://www.mckinsey.com/~/media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/getting%20the%20most%20out%20of%20industry%204%200/mckinsey_industry_40_2016.ashx (Link last accessed 01-16-25.)
4. "German Industry 4.0 Index 2018," Staufen AG; (2018). https://www.staufen.ag/fileadmin/HQ/02-Company/05-Media/2-Studies/STAUFEN.-Study-Industry-4.0-Index-2018-Web-DE-en.pdf
5. "German Industry 4.0 Index 2019," Staufen AG; (2019). https://www.staufen.ag/fileadmin/HQ/02-Company/05-Media/2-Studies/STAUFEN.-Study-Industry-4-0-index-2019-en_.pdf
6. "German Industry 4.0 Index 2022," Staufen. https://en.staufen.ag/insights/studies-and-whitepapers/study-german-industry-4-0-index-2022/ (Link last accessed 01-16-25.)
7. Vuksanović, Herceg, I., Kuč, V., Mijušković, V.M., Herceg, T., "Challenges and Driving Forces for Industry 4.0 Implementation," *Sustainability*, 2020 May 21;12(10):4208. https://www.mdpi.com/2071-1050/12/10/4208 (Link last accessed 01-16-25.)
8. Raj, A., Dwivedi, G., Sharma, A., Lopes de Sousa Jabbour, AB, Rajak, S., "Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: An inter-country comparative perspective," I*nternational Journal of Production Economics.* 2020 Jun;224:107546. https://linkinghub.elsevier.com/retrieve/pii/S092552731930372X (Link last accessed 01-16-25.)

9. Kumar, S., Raut, R.D., Aktas, E., Narkhede, B.E., Gedam, V.V., "Barriers to adoption of industry 4.0 and sustainability: a case study with SMEs," *International Journal of Computer Integrated Manufacturing*, 2023 May 4;36(5):657–77. https://www.tandfonline.com/doi/full/10.1080/0951192X.2022.2128217 (Link last accessed 01-16-25.)

10. Frank AG, Dalenogare LS, Ayala NF., "Industry 4.0 technologies: Implementation patterns in manufacturing companies," *International Journal of Production Economics.* 2019 Apr; 210:15–26. https://linkinghub.elsevier.com/retrieve/pii/S0925527319300040 (Link last accessed 01-16-25.)
11. Zheng, T, Ardolino, M., Bacchetti, A., Perona, M., "The applications of Industry 4.0 technologies in manufacturing context: a systematic literature review," 2021 Mar 19;59(6):1922–54. https://www.tandfonline.com/doi/full/10.1080/00207543.2020.1824085 (Link last accessed 01-16-25.)
12. Landeta Echeberria, A., "A Digital Framework for Industry 4.0: Managing Strategy." Cham: Springer International Publishing; 2020. http://link.springer.com/10.1007/978-3-030-60049-5 (Link last accessed 01-16-25.)
13. Vaidya, S., Ambad, P., Bhosle, S., "Industry 4.0 – A Glimpse," *Procedia Manufacturing,* 2018;20:233–8. https://linkinghub.elsevier.com/retrieve/pii/S2351978918300672 (Link last accessed 01-16-25.)
14. Normando Souza, L., Fonseca, T., Reis Vasconcelos, J., Barbosa Da Cruz, C.A., Almeida Paixão, A.E., "Pillars of Technological Advancement and Indicators of Scientific Production of Industry 4.0," *Int J Innov Educ Res*. 2020 Apr 1;8(4):245–57. https://scholarsjournal.net/index.php/ijier/article/view/2286 (Link last accessed 01-16-25.)
15. Blondheim, Jr., D., "The Promise of Industry 4.0 for North America's Foundries: Building a Connected Foundry," American Foundry Society (AFS) 2023 Leadership Summit; (2023 Sep 19); Asheville, NC.
16. Blondheim, Jr., D., "Die Casting in the Digital Age: Harnessing Industry 4.0 to Build a Connected Foundry," North American Die Casting Association (NADCA) 2024 Executive Conference; (2024 Feb 21); Phoenix, AZ.
17. Blondheim, Jr., D., "A Roadmap to Build a Connected Foundry," Keynote presented at: American Foundry Society (AFS) 2024 Foundry 4.0 Conference; (2024 Jul 30); Milwaukee, WI.
18. Waibel, M.W., Steenkamp, L.P., Moloko, N., Oosthuizen, G.A., "Investigating the Effects of Smart Production Systems on Sustainability Elements," *Procedia Manufacturing,* 2017;8:731–7. https://linkinghub.elsevier.com/retrieve/pii/S2351978917301002 (Link last accessed 01-16-25.)
19. Kayikci, Y., "Sustainability impact of digitization in logistics," *Procedia Manufacturing,* 2018;21:782–9. https://linkinghub.elsevier.com/retrieve/pii/S2351978918302245 (Link last accessed 01-16-25.)
20. Mohamed, M., "Challenges and Benefits of Industry 4.0: An overview," 2018 Aug;5(3). http://www.ijsom.com/article_2767_1ed3957e521af286722efb31b2772314.pdf (Link last accessed 01-16-25.)
21. "The Moving Assembly Line," Ford Corporate. https://corporate.ford.com/articles/history/moving-assembly-line.html (Link last accessed 01-16-25.)
22. Dossou, P.E., "Development of a new framework for implementing industry 4.0 in companies," *Procedia Manufacturing* (2019) 38:573–80. https://linkinghub.elsevier.com/retrieve/pii/S2351978920300731
23. Gajdzik, B., Grabowska, S., Saniuk, S.A., "Theoretical Framework for Industry 4.0 and Its Implementation with Selected Practical Schedules," *Energies* (2021 Feb 10);14(4):940. https://www.mdpi.com/1996-1073/14/4/940 (Link last accessed 01-16-25.)
24. Çınar, Z.M., Zeeshan, Q., Korhan, O.A, "Framework for Industry 4.0 Readiness and Maturity of Smart Manufacturing Enterprises: A Case Study," *Sustainability* (2021 Jun 11);13(12):6659. https://www.mdpi.com/2071-1050/13/12/6659 (Link last accessed 01-16-25.)
25. Protocol - Glossary | CSRC. NIST - Computer Security Resource Center. https://csrc.nist.gov/glossary/term/protocol (Link last accessed 01-16-25.)
26. "Management of alarm systems for the process industries," Research Triangle Park, North Carolina: International Society of Automation; (2016). 978-1-941546-86-4
27. Durana, P., Kral, P., Stehel, V., Lazaroiu, G., Sroka, W., "Quality Culture of Manufacturing Enterprises: A Possible Way to Adaptation to Industry 4.0.," *Social Sciences* (2019 Apr 19);8(4):124. https://www.mdpi.com/2076-0760/8/4/124 (Link last accessed 01-16-25.)
28. "Labor Productivity and Cost Measures: Major sectors: nonfarm business, business, nonfinancial corporate, and manufacturing," (September 5, 2024). https://www.bls.gov/productivity/tables/labor-productivity-major-sectors.xlsx (Link last accessed 01-16-25.)
29. Andrews, Evan,"Who Were the Luddites?" *HISTORY* (8-21-2023). https://www.history.com/news/who-were-the-luddites (Link last accessed 01-16-25.)
30. Stamp, Jimmy, "The Many, Many Designs of the Sewing Machine,"| *Smithsonian* https://www.smithsonianmag.com/arts-culture/the-

many-many-designs-of-the-sewing-machine-2142740/ (Link last accessed 01-16-25.)
31. Rosen, L., Sears, D.C., Wiel, M.M., "COMPUTERPHOBIA," Behavior Reasearch Methods, Instruments, & Computers. 1987;19(2):167–79. https://doi.org/10.3758/BF03203781
32. Mohelska, H., Sokolova, M., "MANAGEMENT APPROACHES FOR INDUSTRY 4.0 – THE ORGANIZATIONAL CULTURE PERSPECTIVE," *Technological and Economic Development of Economy* (2018 Nov 21);24(6):2225–40. https://journals.vilniustech.lt/index.php/TEDE/article/view/6397 (Link last accessed 01-16-25.)