



Industrial Analytics: The Engine Driving the IIoT Revolution

An Industrial Internet Consortium Technical White Paper

Wael William Diab, Huawei

K. Eric Harper, ABB

Dr. Shi-Wan Lin, Thingswise, LLC

William Sobel, System Insights

IIC:WHT:IN2:V1.1:PB:20170329

1 INTRODUCTION

The Industrial Internet of Things (IIoT) seeks to connect industrial assets and machines—the things—to enterprise information systems, business processes and people who operate and use them. Advanced analytics is at the core of this next-generation level of integration and, when applied to machine and process data, provides new insights and intelligence to optimize decision-making significantly and enable intelligent operations leading to transformational business outcomes and social value. If data is the new fuel, data analytics is the new engine that propels the IIoT transformation. As a fledgling new discipline combining advances in mathematics, computer science and engineering in the context of Information Technologies (IT) and Operational Technologies (OT) convergence, industrial analytics plays a crucial role in the success of any IIoT system. Operating in industrial settings, analytics has unique requirements, characteristics and challenges compared to business analytics and so requires special considerations in its implementation. Much needs to be explored since industrial analytics is in its early stage of development. This paper is intended to spur discussions and research, and speed up the development and maturity of this indispensable technology.

2 THE INDUSTRIAL INTERNET OF THINGS

The Industrial Internet of Things is a natural extension of the Industrial and Internet revolutions.

IIoT will be a major force driving economic growth for the coming decades, at a greater pace than prior revolutions. As outlined by the World Economic Forum¹, “The first Industrial Revolution used water and steam power to mechanize production. The Second used electric power to create mass production. The Third used electronics and information technology to automate production. Now, a Fourth Industrial Revolution is building on the Third, the digital revolution that is blurring the lines between the physical, digital and biological spheres.”

To accelerate this digital revolution the Industrial Internet Consortium (IIC) is advancing the technology of IIoT across a diverse set of application domains. The Industrial Internet integrates the industrial assets and machines—the things—to enterprise information systems, business processes and people who operate or use them. With these connections to the industrial assets and machines, new technologies enable the application of advanced analytics to machine and operational process data to gain insights into the operations, optimize them intelligently to boost productivity, increase quality, reduce energy and material consumption, increase flexibility, and ultimately create new business values, while maintaining commitments to safety, reliability, security and data privacy, and conservation of the environment as social values.

The landscape for IIoT encompasses both IT and OT. Industrial analytics, analytics applied to machine data for operational insights, can be seen as an engine driving the convergence of OT and IT, and ultimately value creation for the Fourth Industrial Revolution.

¹ Schwab, K., “The Fourth Industrial Revolution: what it means, how to respond”, World Economic Forum (2016).

3 THE VALUE OF INDUSTRIAL ANALYTICS

Analytics may be broadly defined as a discipline transforming data into information through systematic analysis. Industrial Analytics is the use of analytics in IIoT systems.

To understand the value of industrial analytics, consider this example use case. In an industrial setting, a major cause for unplanned downtime and expense is machine outage. This translates to billions of dollars² lost in unplanned failures of equipment and unnecessary maintenance. Currently, most companies use scheduled maintenance plans for preventive maintenance. This means that maintenance is performed on machines even if they are sitting idle, costing time and resources, as well as requiring unnecessary disruptive procedures that can reduce equipment reliability. On the other hand, critical problems are often missed due to insufficient diagnostics resulting in unplanned downtime and often-costly repairs. Both the over- and under-maintenance of assets contribute to higher operating expense (OpEx).

To address these issues, maintenance needs to shift to prognostics that will only schedule maintenance based on component lifetime characteristics and its usage, rather than solely on schedule. Next, the practice needs to advance to predictive where analytics are applied to sensor and machine operational data to forecast the likelihood of certain failures in a given period. With this information, machine maintenance can be optimally scheduled and performed to avoid unexpected interruption to the operations and at lower cost. Analytics can even be applied into the control loop to adapt the machines to avoid operational conditions that can have detrimental effects on the machines and processes.

Data analytics plays a vital role providing crucial insights needed for decision-making and for optimal deployment of resources in industrial applications. Those capabilities in turn effectively translate to an increase in the efficiency of labor and capital. Long-term GDP growth is fundamentally dependent on capital, labor and efficiency. Increasing the efficiency of capital and labor is paramount to having not only a path for GDP growth but also competitive advantages, whether it is in developed or emerging economies.

4 ANALYTICS IN INDUSTRY

Analytics as a discipline in applied mathematics has been widely adopted by organizations for decades, most prominently as business analysis in finance, banking and ecommerce.

In industrial settings, analytics is currently used to identify and address potential faults in assets, improve uptime and reduce repair costs. This is called “condition monitoring.” With the rapid development of information and communication technology (ICT), partly aided by standards and innovations in sensor and computer technology, it is now possible to extend advanced analytics to large numbers of machines across the globe. Advances in analytics algorithms and techniques,

²J. Manyika, et al, “The Internet of Things: Mapping the Value Beyond the Hype”, McKinsey & Company (June 2015)

including machine learning, are used to analyze large amounts of data gathered from industrial control systems. The insights gained from the analytics can be applied automatically to increase operational efficiency of machines by, for example, anticipating peak usage, streamlining the supply chains for parts needed for preventative maintenance, and for business planning and decision-making. Using insights drawn from the machine data to drive intelligent operational and business processes, industrial analytics enables the convergence of analytics in the OT and IT worlds.

5 DESIGN CONSIDERATIONS

One of the common questions is where the analytics should be performed.

The consideration for industrial analytics location may also come in the form of “do I need analytics at the edge?” or “is it sufficient to have my analytics engine run on top of my big data platform?” The answer is, it depends. Let's take a closer look at how an industrial ecosystem works from the business and ICT perspectives. The following constraints may govern the general operation and associated requirements for industrial analytics.

Scope: Ultimately, it is the derived information (not the raw data) and how it can be acted on that determines what kinds of analytics are deployed. For instance, if the goal is to optimize machine uptime at one site, then analytics performed on data gathered there may be sufficient. In this case, the analytics can be performed anywhere, provided that, if done remotely, the normal local operation is not critically dependent on network latency and the availability of the analytics results. On the other hand, if the value proposition is to optimize production across sites requiring comparison of factory efficiencies then analytics needs to be performed on data gathered from these sites and thus be available in a higher tier in the system architecture.

Response time and reliability: It is expected in an industrial setting that some problems require deterministic analysis, computation and response. Others can be done after the fact. For the former, it almost always requires analytics “locally” for reliability and performance.

Bandwidth: With advancements in sensor technology and ever increasing cost reductions in its deployment, a large number of sensors may be deployed in an industrial setting. The total amount of data generated by these sensors, together with data gathered from the control systems can be massive. The increased network and infrastructure required to fuse the massive amount of data from one domain with those in others will be compensated by the creation of valuable insights.

Capacity: In some cases, it may be optimal to perform the analytics at a particular tier in a system architecture, but the existing infrastructure may not be able to support it, so a different tier is selected. Key properties of the ICT infrastructure include latency, bandwidth and computational capacity.

Security: The value from moving data must be balanced with concerns for transferring raw data outside of controlled areas, and the associated costs. Thus, it may be more efficient to perform

certain analytics locally first and only share necessary summary, redacted or anonymized information with those in the other domains.

Volume: All data needs to be stored, at least temporarily, to derive business value. The storage capacity required depends on the application, but in general there is more available and at a lower cost the closer you get to the cloud.

Velocity: Industrial measurements are typically captured cyclically. Typical high-frequency data, such as vibration or acoustics data from aircraft engines and wind turbines, can significantly increase the speed of data processing. Another consideration is transient events where readings must be captured with accurate time recording to determine order of occurrence, causality and root cause. With these low latency requirements, analytics using high velocity data is better performed close to where the data is measured.

Variety: In many industrial settings, a major impediment to rapid innovation may be the variety of data, not the velocity or volume. Such a setting often comprises many pieces of equipment that have been acquired over many years, having dissimilar controls, interfaces and available data. Effective analytics depends on shared information models to understand the data highly varying in both format (syntax) and content (semantics) to deliver the expected insights.

Analytics maturity: Processing of raw data (measurements) into descriptions (information) and then contextualizing the results (knowledge) and benefiting from historical experience (wisdom) is not limited to where the analytics can be performed.

Temporal Correlation: One of the common issues in IIoT systems is correlating data between multiple sensors and process control states, including the temporal order at which the data is generated. Simply progressing the data to upper tiers without any local analysis may increase the burden of correlation when the analytics is applied. By applying analytics at a lower tier closer to where the data is generated addresses the temporal order issue.

Provenance: As data progresses through an IIoT system into different domains and upwards through the tiers in the system architecture, it is easy to lose track of the genuine source of the data. Performing analytics at a lower architecture tier avoids losing track of the genuine sources of the data as they progress through the IIoT system.

Compliance: As an example, national security concerns may place restriction on the architectural decision about data management and sharing with government regulations in industries like aerospace and defense. This will influence where the analytics must be placed to meet the regulatory requirements, for example, possibly preventing performing large-scale computation in a public cloud facility to lower the cost.

Table 1, below, illustrates example considerations of the feasibility of locating industrial analytics processing in a deployment, organized in this instance as the *plant*, *enterprise* and *cloud*, where the first two are hosted within the organization and the third provided outside but still under their control. For any one consideration, there is flexibility for the location, but considering more than one quickly narrows the feasible alternatives.

Industrial Analytics Location

Evaluation Criteria	Plant	Enterprise	Cloud
Analysis Scope			
Single site optimization	X	X	X
Multi-site comparison		X	X
Multi-customer benchmarking			X
Results Response Time			
Control loop	X		
Human decision	X	X	
Planning horizon	X	X	X
Connectivity Reliability			
Site	X		
Organization	X	X	
Global	X	X	X
Connectivity Bandwidth			
Raw data	X		
Processed results	X	X	
Summarized results	X	X	X
Storage and Compute Capacity			
Server	X	X	X
Multiple servers		X	X
Data center			X
Data Security			
Secret	X		
Proprietary	X	X	

Industrial Analytics Location

Evaluation Criteria	Plant	Enterprise	Cloud
<i>Shared</i>	X	X	X
Data Characteristics			
<i>Volume</i>			X
<i>Velocity</i>	X		
<i>Variety</i>	X	X	X
Analytics Maturity			
<i>Descriptive</i>	X	X	X
<i>Predictive</i>	X	X	X
<i>Prescriptive</i>	X	X	X
Event Correlation			
<i>Sub-seconds</i>	X		
<i>Seconds</i>	X	X	
<i>Tens of seconds</i>	X	X	X
Data Provenance			
<i>Sensor</i>	X		
<i>Asset</i>	X	X	
<i>Site</i>	X	X	X
Regulatory Compliance			
<i>Asset</i>	X	X	X
<i>Process</i>		X	X
<i>Industry</i>			X

X = Feasible alternative

Table 1. Industrial Analytics Design Considerations

It is the combination of all these factors that determines what capabilities are required and where the analytics will be deployed. Generally speaking, which analytics and where the processing is located depends on the maximum acceptable network latency and jitter in response to events, criticality of the analytics in the normal operations (e.g. how bad it would be if the external network is cut or an upper tier system becomes unavailable) and the cost of uploading large amounts of data. From the deterministic response, reliability and resilience perspectives, it is optimal to perform analytics close to the sources of the data, and the analytics results need to be accessible for decision-making. In most systems, some form of hybrid approach with local and centralized analytics will be required.

6 CREATING BUSINESS VALUE

An enterprise needs to increase throughput, reduce expenses and inventory to generate higher margins.

Sales drives manufacturing throughput, but it cannot exceed the enterprise's capacity to produce, otherwise compromising on-time delivery and disappointing customers. One approach is to identify performance bottlenecks in overall operations continuously and remove them one-by-one to meet the demand from sales and profit targets. This process needs to be grounded in reliable information about the demands, production, inventory and the operational processes. IIoT and specifically industrial analytics enables stakeholders to gather data from machines and optimize processes more efficiently. The opportunity is substantial since over 86% of companies do not have a corporate analytics program that uses manufacturing data (LNS Research, 2016³).

According to the 2016 Global Manufacturing Competitiveness Index report by Deloitte Touche Tohmatsu Limited and the US Council on Competitiveness, Predictive Analytics and Smart Connected Products are #1 and #2, respectively (see Table 2, on the following page). This shows that the desire to leverage analytics is a high priority for most CEOs, but the awareness of how that technology will be applied and the areas of greatest impact are not well understood.

³ <http://blog.insresearch.com/5-real-surprises-from-the-2016-metrics-that-matter-research-study>

Advanced Manufacturing Technologies	US	China	Europe
Predictive analytics	1	1	4
Smart, connected products (IoT)	2	7	2
Advanced materials	3	4	5
Smart factories (IoT)	4	2	1
Digital design, simulation, and integration	5	5	3
High performance computing	6	3	7
Advanced robotics	7	8	6
Additive manufacturing (3D printing)	8	11	9
Open-source design/Direct customer input	9	10	10
Augmented reality (to improve quality, training, expert knowledge)	10	6	8
Augmented reality (to increase customer service & experience)	11	9	11

Table 2. Ranking of future importance of advanced manufacturing technologies by executives, Source: Deloitte Touche Tohmatsu Limited and US Council on Competitiveness, 2016 Global Manufacturing Competitiveness Index

As our business systems and analytical capabilities increase, some enterprises will adopt data and analytical strategies, some will not. Enterprises that avail themselves of these technologies will far outpace the others in their ability to produce products and services quickly and safely and maintain quality with full accountability for their processes. The cost of manufacturing will also be flattened, allowing developed countries to compete effectively in global markets because the value of skills and experience will more than compensate for the wage differential.

7 ANALYTICS ARCHITECTURE

Industrial analytics can be applied in different domains distributed across an IIoT system and tuned to varying time scale horizons.

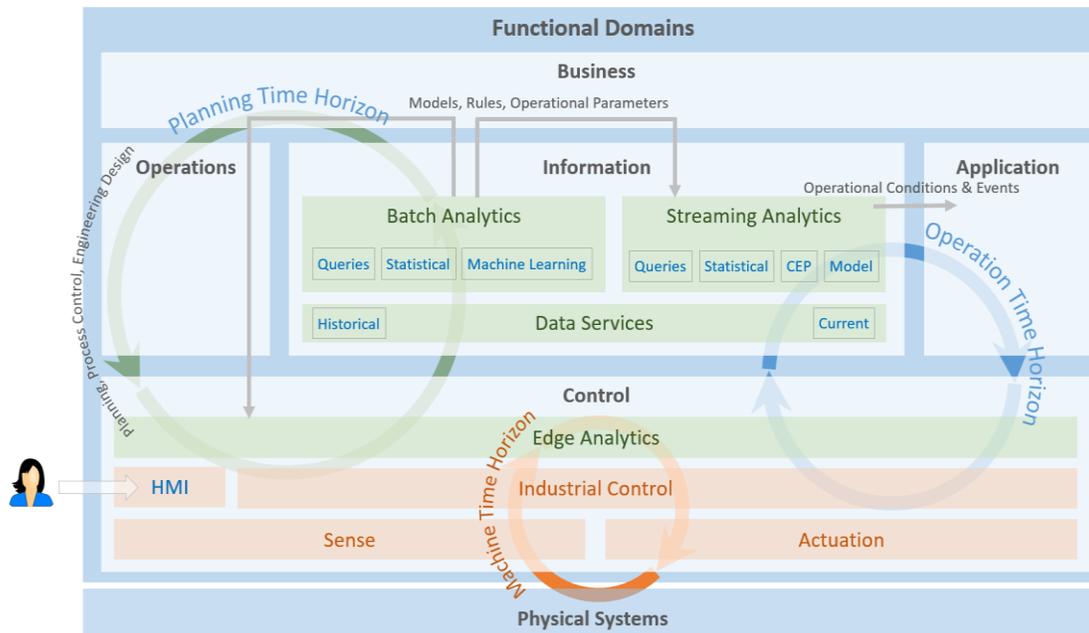


Figure 1. Analytics Mapping to the Industrial Internet Reference Architecture

The required architecture can be analyzed in the context of the functional domains of the Industrial Internet Reference Architecture (IIRA) developed by the IIC, as shown in Figure 1 above. An end-to-end IIoT system in the IIRA is functionally decomposed into five functional domains:

1. **Control:** sensing, communication, execution, motion, and actuation;
2. **Operations:** provisioning, management, monitoring, diagnostics and optimization;
3. **Information:** data fusion, transforming, persisting, modeling and analyzing;
4. **Application:** logic, rules, integration, human interface; and
5. **Business:** enterprise and human resources, customer relationships, assets, service lifecycle, billing and payment, work planning and scheduling.

The control domain is a collection of functions that are performed by the industrial assets or control systems, most predominantly, in exercising the closed control loops within themselves. The operations domain is a collection of functions for assets and control systems management and maintenance to ensure their continuing operations. The Information domain is a collection of functions for collecting, transforming, analyzing data to acquire high-level intelligence of the entire system. The Application domain is a collection of functions for applying use case specific logic, rules and models based on the information obtained from the information domain to achieve system-wide optimization of operations or other business objectives. The business domain is a collection of functions for integrating information across business systems and applications to achieve business objectives.

Industrial analytics results can be applied to the control domain as edge analytics providing real-time operational insights to the control loops in a *machine time horizon* that typically requires analytic response in milliseconds or less. Examples include autonomous vehicles and robotics. Analytics in this time horizon tends to be streaming in nature and applied automatically. Industrial analytics results can be applied to the application and operations domains to provide machine insights that enable advanced maintenance such as automatic fault detection and diagnosis, and preventive maintenance, or to drive optimal operations across fleets of machines or assets. The analytics result is applied in an *operation time horizon* that typically requires an analytic response in the range of seconds or more. Analytics in this time horizon also tends to be streaming in nature and applied automatically. Industrial analytics results can also be applied to the business domain as well, providing insights to enable intelligent business processes, including aiding business planning and engineering processes. The analytics result is applied in a *planning time horizon* that typically requires an analytic response in the range of days or more. It consists of both streaming analytics results that are applied automatically (e.g. to work and machine part scheduling for on-site repair) and batch analytics results based on on-demand queries.

8 GETTING STARTED WITH INDUSTRIAL ANALYTICS

Industrial analytics are used to identify and recognize machine operational and behavioral patterns, make fast and accurate predictions and act with confidence at the points of decision.

Analytics generally fall into three major categories:

Descriptive analytics gain insight from historical or current data streams including for status and usage monitoring dashboard, reporting, anomaly detection and diagnosis, model building or training and such.

Predictive analytics identify expected behaviors or outcomes based on predictive modeling using statistical and machine learning techniques, e.g. capacity demand/usage prediction, material/energy consumption prediction, and component/system wear and fault predictions.

Prescriptive analytics finds the optimal solution by determining what is likely to happen based on first principles, empirical models and using predictive analytics that incorporates causality related to design and execution decisions. An example of prescriptive analytics is on-demand production from a solid geometric assembly model to find the optimal set of manufacturing processes to achieve the final product intent, taking into consideration all possible options and capabilities.

The analytics results can be applied automatically to the machines and systems, or used to support human decisions through visual analytics to enhance human understanding and generate confidence in a decision.

Industrial analytics has unique challenges as an engine driving IIoT transformation because the results can alter the operation and safety of things in the physical world. These effects may be undesirable or harmful, inadvertently affecting the safety of people or damaging property and the environment. Moreover, because industrial analytics often interpret data from different sensors and machines that may conflict with one another, we need to understand and synthesize

the diverse information streams to reach a correct conclusion. Table 3, below, identifies the unique requirements to consider when planning for industrial analytics.

<i>Correctness</i>	Industrial Analytics must satisfy a higher level of accuracy in its analytic results. Any system that interprets and acts on the results must have safeguards against undesirable and unintended physical consequence.
<i>Timing</i>	Industrial Analytics must satisfy certain hard deadline and synchronization requirements. Near instantaneous analytic results delivered within a deterministic time window are required for reliable and high quality actions in industrial operations.
<i>Safety</i>	When applying Industrial Analytics, and interpreting and acting on the result, strong safety requirements must be in place safeguarding the wellbeing of the workers, users and the environment.
<i>Contextualized</i>	The analysis of data within an industrial system is never done without the context in which the activity and observations occur. One cannot construct meaning unless a full understanding of the process that is being executed and the states of all the equipment and its peripherals are considered to derive the true meaning of the data and create actionable information.
<i>Causal-oriented</i>	Industrial operations deal with the physical world and Industrial Analytics needs to be validated with domain-specific subject matter expertise to model the complex and causal relationships in the data. The combination of first principles, e.g. physical modeling, along with other data science statistical and machine learning capabilities, is required in many industrial use cases in order to provide accurate analytics results.
<i>Distributed</i>	Many complex industrial systems have hierarchical tiers distributed across geographic areas. Each of these subsystems may have unique analytic requirements to support their operations. Therefore, Industrial Analytics must be tailored to meet the local requirements of the subsystems it supports. The requirements on timing (avoiding long latency) and resilience (avoiding widespread outage of service because of faults in the network or in a centralized system) require a distributed pattern of Industrial Analytics in that the analytic will be implemented close to the source of data it analyzes and to the target where its analytic outcome is needed.
<i>Streaming</i>	Industrial Analytics can be continuous or batch processes. Because of continuous execution in industrial systems, a large proportion of Industrial Analytics will be streaming in nature, performing analysis of live data and providing continuous flow of analytics results in support of the operations. Traditional batch-oriented analytics will still be performed either for building or improving analytic models, or for human decision-making.
<i>Automatic</i>	In order for the Industrial Analytics to support continuous operations, the analysis of streaming data and the application of analytic outcomes must be automatic, dynamic and continuous. As the technologies in Industrial Analytics advance, improvements in analytic modeling e.g. through learning may also be automatic.
<i>Semantics</i>	Analytical systems require data that has meaning and context. Unstructured data, when reported without attribution to the source and the component or system it represents, makes deriving value complex since it requires the analytics to guess or infer the meaning. Inference unnecessary adds significant uncertainty into the system. Most data can be properly attributed at the source, and if this information is communicated, it can significantly increase the success and accuracy of the analytical systems.

Table 3. Industrial Analytics Requirements

9 ANALYTICS CAPACITY CONSIDERATIONS

The functionality for industrial analytics is informed by the experience and wisdom from both IT and OT. Both expect reliable operation and repeatable response times. However, the way these goals are achieved is different. Information technology relies on elasticity to provide the required capacity and operational technology ensures determinism with engineered capacity.

Elasticity is a fundamental cloud technology measure of “the degree to which a system is able to adapt to workload changes by provisioning and de-provisioning resources autonomic manner, such that at each point in time the available resources match the current demand as closely as possible.”⁴ For example, most retail businesses generate significant revenue during the holiday season at the end of the year. Dramatic changes in staffing, production and logistics are made in preparation for this crucial time-period. Correspondingly, all the IT systems need to be prepared and have enough capacity to handle these changes to avoid affecting profitability. The remainder of the year those resources can be scaled back or repurposed to reduce costs.

Determinism is the ability to support predictable computation and transmission of measurements and results between the devices attached to the network. Deadlines must be met with expectations that a work request will complete with the same response time for every request. The analytics and results must be communicated in a defined time-period and confirmation must be provided. For example, operational systems are designed with a capacity for continual processing, regardless of the state or condition of the plant. So, when a production plant starts up or shuts down, which can generate rapidly changing values and multiple alarms, the requirement for deterministic response in the IIoT system is no different than when the plant is at steady state.

These two philosophies are in many ways complementary. Close to industrial processes, it is appropriate to provide reliability and predictability using dedicated resources. In the cloud with multiple tenants sharing the same set of resources dynamically, it makes more sense to deliver capabilities just-in-time because there could be different service levels. There are early signs that the manufacturing industry is moving in the direction of on-demand services to deliver just-in-time parts using shared assets. The World Economic Forum, in collaboration with Accenture, have predicted that this will become the long-term trend of manufacturing and there will be a convergence between the concerns between shared IT capacity on demand and manufacturing services.⁵ In other industries, such as transportation and energy, there is movement towards services and shared capacity.

⁴ Herbst, Nikolas Roman; Samuel Kounev; Ralf Reussner (2012). "Elasticity in Cloud Computing: What It Is, and What It Is Not" (PDF). Proceedings of the 10th International Conference on Autonomic Computing (ICAC 2013), San Jose, CA, June 24–28.

⁵ O’Halloran, et. al., (2015). “Industrial Internet of Things: Unleashing the Potential of Connected Products and Services” (PDF). World Economic Forum in collaboration with Accenture.

10 ANALYTICS FUNCTIONALITY

Industrial Analytics functionality is deployable throughout the architecture.

The capabilities needed for successful industrial analytics solutions are shown in Table 4, below. Each feature is realized by a set of functions defined by use cases that meet the stakeholders' expectations, especially with regard to non-functional requirements.

<i>Visualize</i>	Display and manage data readings and analytics results using a common framework
<i>Explore</i>	Ad-hoc experiments with historical data
<i>Design</i>	Automation of the data analytics stages; data quality, data mining, and business intelligence algorithm composition
<i>Orchestrate</i>	Delegate work requests over a cluster of computing resources, and collect and aggregate intermediate and final results
<i>Connect</i>	Exchange data and work requests between components using a common framework
<i>Cleanse</i>	Merge data sets from different data sources based on suitable criteria; remove irrelevant data and clean noise from data
<i>Compute</i>	Perform computation of statistical, first principle and machine learning model analytical calculations, including live analysis on streaming data, batch or ad hoc data mining and operation and business intelligence analysis
<i>Validate</i>	Ensure analytics results when applied in the context of the application and environment will not harm people or property. This function should be independent from the core analytics processing and act as a governor
<i>Apply</i>	Apply analytics results to various subsystems, including the automation systems (e.g. adjusting control parameters or models), operations and business processes, increasingly automatically or as information aiding human decision-making
<i>Store</i>	Archive and reproduce measured and calculated data streams, especially time-series sequences
<i>Manage</i>	Manage the information model, including data sources, computing resources and data analytics metadata
<i>Supervise</i>	Manage system reliability by ensuring processes are started and maintained, and that computer resources are not exhausted

Table 4. Industrial Analytics Capabilities

For example, in our manufacturing use case, the plant operators visualize the conditions of the process line using graphical displays. When an alert appears, the operator uses the display to drill down into the time-series records of the key parameters. On the other hand, in planning for the next run, the operator might use a similar drill-down to visualize the workflow and determine if there are any concerns that need to be addressed based on anticipated environmental conditions.

The fundamental prerequisite for industrial analytics is availability and access to data from the industrial process and related assets. Data is collected close to the process through connections and stored, at least temporarily, where the readings can be scanned and evaluated depending on the type of analytics. The stored values may be discarded or archived for further calculations. Data scientists can explore the archived data using statistics to compute correlations, and apply algorithms to classify and cluster the evidence over time. Industry subject matter experts have a good understanding of the context and condition of the process and assets, and can interpret and validate the readings and recommend cleansing filters. It is this combination of data science and subject matter expertise that produces the best results.

The industrial analytics workflow can be automated as the data and relationships are better understood. Given an appropriate framework, the automation is designed, configured and orchestrated to transform raw data into actionable results. The workflow and algorithmic content is versioned and deployed both on premise and in the cloud as needed to meet stakeholder expectations. The entire process is supervised to ensure all steps are completed and validated. Ideally, the industrial analytics solution evolves over time to produce better results using a minimum of resources, and to improve accuracy as more experience and historical data are acquired.

The final step is to communicate and present the industrial analytics results in a compelling and easily understandable format. This functionality many times is unique to the application, but can employ a framework for display, including charts, graphs, and recommended actions. Most important is to provide means for humans to interact with the results, starting with a summary and allowing drill down into the evidence that supports the recommendations.

As analytics advances, more and more meaningful operational patterns, especially anomaly, will be detected, identified and reported as alerts, along with relevant supporting data to the operators automatically. Moreover, root causes of faults can be automatically diagnosed and remedies or repairing actions can be prescribed; faults and failures can be prevented by ruling out improper operating parameters that are outside of the normal ranges and be predicted based on the historical experience of making similar parts. On the other hand, machine operational efficiency can be monitored and optimized based on analytics result as well as orchestration of manufacturing resources and coordination and interaction of the prices of equipment and the people who operate them.

Clearly, all these will improve the operational efficiency of manufacturing and operations and at the same time reduce the stress on human operators in keeping the machines operating at their best capacity. On the other hand, analytics is no magic by itself—it requires a combination of obtaining the proper data at the proper time, applying the proper analytics algorithms and

models which are guided by the necessary engineering domain knowledge from both the machine manufacturers, system integrators and the plant operators themselves. It is important to recognize that this is a process of continuing learning and improvement as any intelligent processes. For example, the types of data collected can be extended and their quality can be enhanced, the algorithms and models can be subjected to refinement and deeper domain knowledge can be injected into the models.

11 CLOSING THOUGHTS

As the engine driving value-creation in IIoT, industrial analytics deserves a great deal of attention from the IIoT community. As a large percentage of industrial companies have not incorporated machine data in their analytics process for decision support and intelligent operations, industrial analytics and IIoT offer a great opportunity to drive the next round of value creation. The development of industrial analytics in an IIoT system needs to be driven by business objectives and value. Industrial analytics, especially at its early stage of development, requires a systematic and architectural approach that is flexible to meet the operational and business requirements and forward looking in accommodating usage changes and technological advances.

Furthermore, to implement analytics that can fully support end-to-end intelligent industrial operations and business processes, it requires cross-service and cross-domain sharing of data and information from analytics. In other words, for example, for manufacturing of products, there are services needed to support the analytical modeling of the design and engineering functions as well as the execution, verification, quality and lifecycle maintenance of that product. These services need to interact to provide optimal solutions by verifying prescriptive decisions with empirical outcome and refining models of both production and design. The interaction of these services and the information sharing will be key to a successful end-to-end analytical solution with greatest impact.

Therefore, the realization of the benefit from industrial analytics requires community and ecosystem-wide efforts in sharing, inventing and standardizing technologies and know-how. With its central role in IIoT and intelligent industrial operations and the great impact it may bring, industrial analytics is expected to experience rapid growth. Because it is still in a nascent stage, it needs collective nurturing for its health development.

Now is the time to build a thriving industrial analytics community to collaborate to achieve its promise.

12 CONTRIBUTOR ACKNOWLEDGMENTS

This document is a work product of the Industrial Internet Consortium. The Industrial Internet Consortium wishes to thank the following contributors for their leadership on this white paper.

Wael William Diab

Wael is a business and technology strategist with more than 860 patents to his name in the fields of networking and ICT. Skilled in strategy for breakout technology products at large companies, Wael has architected strategies, driven industry standards and led associated company technology roadmaps for Green Networking, Ethernet, Power over Ethernet, Ethernet in the First Mile (EFM), Cloud/Mist/Fog/Edge Networking, Internet of Things (IoT), Industrial Internet, Big Data, automotive networking, Intelligent Transportation Systems (ITS) and audio video bridging, Time Sensitive and Deterministic Network, ICT Sustainability, Open Compute and open source, amongst others.



Wael is currently a Senior Director at Huawei. Prior to Huawei, Wael held senior leadership roles at Fortune 500 companies, including Cisco and Broadcom. Wael holds BS and MS degrees in Electrical Engineering and a BA in Economics from Stanford, and an MBA with honors from the Wharton School of Business. He has 358 US issued patents and over 504 pending US patent applications in the ICT field. Wael was awarded the TechAmerica David Packard 2011 Medal for innovation in green technology. Wael co-chairs the IIC Liaison Working Group, Technology Working Group, Industrial Analytics Task Group and is a member of the IIC Steering Committee as Huawei's alternate. Wael is a published author, having coauthored *Ethernet in the First Mile: Access for Everyone*.

K. Eric Harper

Eric is a senior principal scientist at ABB Corporate Research specializing in software architecture and sustainable technologies. Over the past nine years he has made significant contributions to the Industrial Internet, process automation, substation automation, and power electronics. Eric graduated from Virginia Tech and attended Rensselaer Polytechnic Institute in electric power engineering as part of the General Electric Advanced Course in Engineering (ABC), working in GE Large Steam Turbine and Corporate Research. Eric transferred to the Westinghouse Power Generation division making substantial technical contributions for a service bureau dedicated to monitoring and diagnosing conditions in electric power plants. Since then Eric has had a variety of assignments over the past two decades as a hands-on project lead, development manager and software architect in distributed architecture, factory automation, business applications, middleware tools, server management, IT security, healthcare IT, software performance and scalability. Eric's current work is to lead ABB's architecture for Industrial Analytics. He served on the Steering Committee for the IEEE/IFIP Working Conference on Software Architecture (WICSA) and represents ABB at the Industrial Internet Consortium as Steering Committee member and Technology Working Group co-chair.



DR. SHI-WAN LIN

Dr. Lin is the CEO and a co-founder of Thingswise, LLC (www.thingswise.com), a startup providing a streaming analytic platform as a turnkey solution that is purposely built for IIoT systems by adapting and innovating key IT technologies to meet OT requirements. Dr. Lin co-chairs various technical groups for the Industrial Internet Consortium, the Architecture Joint Task Group between Plattform Industrie 4.0 and the Industrial Internet Consortium and the National Institute for Standards & Technology (NIST) Cyber-Physical Systems Public Working Group. Dr. Lin is a leading contributor to the Industrial Internet Reference Architecture published by IIC. Previously, he worked for Intel for 10 years, last as a Principal Engineer/Technologist in the Strategy and Technology Office in its Internet of Things Group. His other prior experience includes working for Sarvega, Inc (a Web Service/SOA/Security startup), Lucent Technologies and Motorola. Dr. Shi-Wan Lin has 20+ years' experience in system architecture, Big Data, analytics, enterprise software, Cloud Service, system security and trust, telecommunications and wireless data communications in both large corporate and startup settings.



WILL SOBEL

Mr. William Sobel is Chief Strategy Officer and co-founder of System Insights and the principal architect and chair of the MTConnect Institute Technical Steering Committee. He also has been the primary contributor to the open source implementations of the MTConnect standard.

Mr. Sobel brings over 25 years of experience architecting, managing, and developing complex applications for numerous industries. Prior to working for System Insights, Mr. Sobel was a visiting lecturer at UC Berkeley working at the RadLAB teaching agile web development. During his employment at the university, Mr. Sobel authored the MTConnect standard as a consultant to AMT (Association for Manufacturing Technology) to address the lack of standardized inter-device communication in manufacturing. Prior to his work at UC Berkeley, Mr. Sobel was VP/Chief Architect at MSCI-Barra, specializing in financial risk management software for some of the largest asset management companies in the world.



A publication of the Technology Working Group, authored by Wael William Diab (Huawei), K. Eric Harper (ABB Technologies), Dr. Shi-Wan Lin (Thingswise), Will Sobel (System Insights).

Copyright © 2017 Industrial Internet Consortium. All rights reserved. This document is provided AS-IS and WITHOUT WARRANTIES.

All copying, distribution and use are subject to the limited License, Permission, Disclaimer and other terms stated in the Industrial Internet Consortium Use of Information – Terms, Conditions & Notices, as posted at https://www.iiconsortium.org/legal/index.htm#use_info. If you do not accept these Terms, you are not permitted to use the document