Technology Report: Data Analytics—From Insights to Actions

We all recognize the value of the data our machines are producing. Many of us are collecting it. And some of us are properly analyzing it. But there remains a gap between these groups and the front-runners who are aggressively using their machine data…transforming insights into actions, relying on data analytics to inform their business strategies and optimize their operations. Knowledge is power only when you do something with it.
CONTENTS

Data analytics for everyone: moving beyond spreadsheets

A pairing of geniuses

The power of real-time, data-driven manufacturing

IIoT data is useless, but the insights are priceless

How tech & data are changing field work

Digital twins for the asset operator
AUTOMATE OPERATIONS. PREDICT QUALITY.

AI for Industrial IoT

**Canvass Analytics** is empowering industrial companies to harness the value of OT and IT data. Using artificial intelligence, the Canvass AI Platform delivers predictive insights to enable the performance, scalability and automation that’s required by today’s smart factory.

**Talk to us today** about improving the quality of our production processes, optimizing energy, and increasing asset performance.
Data analytics for everyone: moving beyond spreadsheets

By Michael Risse, vice president, Seeq Corporation

Love it or hate it, the spreadsheet has been the backbone of every engineer’s analytics effort for close to 40 years. With spreadsheets approaching middle-age, now older than many of their users, there is a sense that we are nearing the end of an era due to the blunt approach of the spreadsheet and associated issues, and due to a new generation of data analytics offerings.

Consider the recent introductions of Schneider Electric’s EcoStruxture products, Rockwell’s Scio and ABB’s Ability platform—with their cloud, advanced analytics and machine learning messaging as common elements. Certainly, these companies’ marketing teams are working hard to brand and promote their new offerings, but at the same time it’s hard to miss the need for vendors to say and deliver something new to the market.

The previous analytics generation, call it the “actionable insights era,” was defined by the spreadsheet and associated applications. It assumed a single, desktop-based user, interacting with disconnected data to produce—following hours of work and labor—a moment of insight.

It’s a condemnation of the actionable insight generation that after so many years of promises, there was so much concern about Grey2K. Where was all the insight and expertise delivered by actionable insights? It was in the heads of the employees streaming out the door, inaccessible at worst and undecipherable at best, in the formulas and files on the hard drives of the departed.

But as we look forward to a new generation of analytics, leaving behind actionable insights, it’s important to recognize that whatever succeeds it has to equal and improve on the core values of the spreadsheet generation. The reality is no product comes close to spreadsheets for data analytics in terms of breadth of users across industries. Despite the fact that spreadsheets seem particularly ill-fitted for working with continuous analog signals, and despite marketing claims of new and improved offerings, its use continues, for two main reasons.

First is simply the accessibility of spreadsheets to users. In an age of technology innovations for other audiences—machine learning for data scientists, cloud computing for IT departments, Google search for consumers—the spreadsheet remains the accessible solution at hand for every engineer.

The second is flexibility: the ability to solve any type of problem, given enough time and effort. If the spreadsheet doesn’t have the needed feature, there’s always a function or code snippet to close the gap, either available online or custom coded.

These foundation points for broad acceptance of spreadsheets must be considered the starting point for any discussion of process industry data analytics going forward. From there, we can consider the following four requirements for what will replace the spreadsheet and address its shortcomings:

- Subject Matter Expert (SME) focus and flexibility
- Simple sharing and saving of analysis results
- Big data and cognitive computing capability
- Process-manufacturing focus
SME FOCUS AND FLEXIBILITY

The first requirement for the next generation of analytics tools is the strength of the spreadsheet user experience, it’s personalization. It turns out that employees who don’t know what they are looking for don’t know what to find, so a self-service, SME focus is the absolute starting point for any new analytics solutions. This requires a reasonable price point, ease of use and a fit for data analytics in process manufacturing.

Meeting these requirements enables data experimentation, enabling SMEs to try things and investigate data, assets and processes without a five-figure purchase requisition or the need for IT assistance. Self-service data analytics delivers ad hoc, ask anything, interaction by SMEs with data located in one or multiple silos.

For SMEs, a self-service approach and flexibility to solve a variety of problems are the nucleus of the frontline analytics experience in process manufacturing.

Spreadsheets have served engineers in process manufacturing for decades, but the time has come for a new and more effective approach to data analytics.
New solutions will offer features paralleling the changes already seen in alternatives for spreadsheets in IT departments, such as cloud-based, multi-user, large scale and visualization solutions from companies such as Smartsheet, Anaplan and Tableau.

It’s interesting that there are so many billion-dollar companies already competing in the IT space as spreadsheet replacements, whereas options for process manufacturing data analytics have lagged behind. This situation is not likely to continue, as data analytics solutions for process industry applications are emerging.

**SIMPLE SHARING AND SAVING OF RESULTS**

As previously discussed, one of the main failures of spreadsheets has been their focus on a single user. New data analytic solutions must therefore move beyond individual users to support collaboration among teams, colleagues and even off-site vendors and partners. These solutions must also support knowledge capture in the analytics process to support greater transparency and re-use of efforts.

The prior spreadsheet model is single user-centric, with collaboration after the fact in network files and SharePoint sites at best, and frequently no sharing of results at worst. Going forward, new data analytics offerings will have collaboration features to enable teams and organizations to participate in the development of solutions, and then easily find and leverage work already completed.

The actionable insights generation had two levels of disconnect. First, data was disconnected from its source and copied or transformed into a spreadsheet for processing. Second, the output of the analytics effort was typically distributed as an image of the output as a static graphic, on paper or in a PowerPoint slide.

In the future, by leveraging a web-based approach to data analytics distribution, a viewer should be able to click through from data visualizations to the underlying data, and even back to the source data. This doesn’t just mean a live dashboard with constantly updating sensor data, it also means automated support for calculated outputs on data streams and batches beyond simple measures and metrics. The result is a single model spanning KPIs and documents that can be easily shifted across time ranges because the underlying connection to the source data is always maintained.

The term sometimes used for these last two points on the engagement and access to insights across the analytics process and outcomes is “data democratisation.” The result is more employees participating, accessing and interpreting data.

**BIG DATA AND COGNITIVE COMPUTING CAPABILITY**

Data volumes, data diversity and data contextualization are the obvious starting point for what the next generation of analytics needs to address. Current volumes are overwhelming existing tools that were never meant for such tasks: millions of rows, thousands of columns. And we should expect only more sensors and more data as wireless data gathering, cloud-based data storage and other technologies continue to advance. The unfortunate result for many process industry firms is lots of data, but not much in the way of useful information.

In particular, the need for data contextualization, ensuring relevant manufacturing and business information is accessible to place data in the wider context of the organization, will grow as more data becomes available from more sources. Typical solutions today may tap up to 10 sources for recipes, sensor data, quality attributes, personnel information, asset data, etc.—each stored in their own silos. The number of data sources required to solve a problem is likely to grow as SMEs are empowered to pursue ad hoc investigations across any data set at any time.

The second component of the next generation for analytics is leveraging the innovation in data management and processing, and cognitive computing.
Saying goodbye to spreadsheets for process data analytics will help firms deliver on the expectations and requirements of new business imperatives.

The more popular terms for these innovations include big data and machine learning, among others. These innovations represent perhaps the biggest gap between the past and future of data analytics, because tapping the incredible intellectual property contributed by the world’s leading software companies will result in radical improvements to the data analytics user experience. Any solution which does not leverage these technologies, and other new advances as they occur, will quickly become obsolete.

PROCESS MANUFACTURING FOCUS

Data cleansing, wrangling and transformation requirements are key reasons why the actionable insights era needs to be replaced with solutions purpose-built for time series data analytics. About 70% of the top 10 Excel functions are for manipulating and organizing data including IF, LOOKUP, VLOOKUP, MATCH, CHOOSE, FIND, INDEX, etc. These functions are required to reformat time series data into rows and columns before the actual insights begin, and thus cut into productivity.

Further, data manipulation activities typically represent 50-90% of the time spent in spreadsheets. These querying, slicing, dicing, sorting, affiliating, synchronizing, interpolating, pivoting, contextualizing and indexing activities must all take place prior to actual work. A new generation of analytics address these mundane, repeated efforts required for working with time series data is long overdue.

Finally, the next generation of analytics solutions must leverage, but not require, a web-based computing model. This will support the current user experience, and a timeline for a gradual adoption of web or cloud deployment aligned to the customer’s timeline.

For the user, this means browser-based access to data analytics from a variety of locations and devices. Engineers shouldn’t have to be where their data or computer is for insights and activity. At the same time, the migration to the cloud for data storage and access will continue to evolve for at least another decade, and in some cases far longer, so a cloud-only deployment model will not suffice for most firms today.

These requirements may be balanced by having a web-based architecture for on-premise or cloud-based deployment, with the ability to connect to data sources residing either on-premise or in the cloud.

CONCLUSION

Saying goodbye to spreadsheets for process data analytics will help firms deliver on the expectations and requirements of new business imperatives such as digital transformation, Industry 4.0 and smart manufacturing. Whatever the term used, the common outcome of these initiatives is delivering benefits to organizations in terms of improved margins, yield, quality and safety by producing insights from new and collected data sources. These benefits will be delivered by a new generation of data analytics offerings.
A pairing of geniuses

By Humera Malik, CEO and founder of Canvass Analytics

Industrial processes are massively complex and generate massive amounts of data. Anyone who has dug their hands into a manufacturing project can attest to that. Industrial operations folks—I consider them marvels—have traditionally used their genius to run these complex networks, optimizing and improving processes on the basis of their experience. But digital transformation and Industry 4.0 are bringing modern approaches to data analysis. And the enterprises that best marry this old-school intelligence with that of the artificial kind (AI) will end up looking the smartest.

AI can be applied to industrial processes for handling the millions of data points being generated by manufacturing environments and predicting future outcomes causing less waste, improved quality and asset optimization for the welding robotic itself.

Whether we’re talking about enhanced quality control or improvements with yield or reductions in process time, data analysis can be the differentiator. And that’s where AI can supercharge outcomes.

AI TRANSFORMING PRODUCTION PROCESSES

If you’re wondering where to get started, AI is best applied to areas with repeatedness, where an automation can be instituted and standards can be set. In situations where you’d previously had an operator managing a process, you can leverage AI (which is not reliant on operator changes) to tackle the complex series of steps in order to reach the desired outcome.

Modern AI is used to create predictions that are constantly learning from the process data.

Efficiency and cost is where most businesses start the process. (Optimizing production is often the following step.) Naturally, enterprises don’t always know what their greatest problem is. The devil is in the details, right? It’s all about asking questions and helping uncover opportunities in which there’s a real need (e.g. problem statement) for optimization, there’s good data, and there’s the right support internally.
Optimize Your Manufacturing Operation, Using Real-Time Production Data

TrakSYS® is a next-generation Manufacturing Operations Management (MOM) software platform designed to help you optimize operations using real-time production data. And to help you take action on those data, TrakSYS includes tools — like alerts, tasks, journals, and more.

TrakSYS: A Standards-Based MOM Platform
With a library of configurable features that conform to the MOM standard architecture, you can use TrakSYS to build solutions for virtually any challenge — from OEE and SPC, to E-Records, Traceability, Workflow, and more. Use TrakSYS to:

- Minimize downtime, shorten changeovers
- Improve quality, reduce errors
- Reduce maintenance time and costs
- Manage inventory, reduce waste
- Comply with regulations and audits

Trusted by Manufacturers Worldwide
First introduced in 1993, TrakSYS is now in use at thousands of factories in over 100 countries, including many of the top brands in pharmaceuticals, food and beverage, packaged goods, automotive, chemicals, and more. Visit us online today to learn how TrakSYS can help you optimize your manufacturing operation, using real-time production data.

Learn More. Watch the Video.
www.parsec-corp.com

Go Further, Go Faster, Go TrakSYS®

©2018 Parsec Automation Corp.

Gain Visibility, Knowledge, and Control
TrakSYS provides a real-time window into production, and gives you the tools to take action on your insights.

Read the Case Study
Learn how Bush Bros. uses real-time production data to optimize its manufacturing operations.

www.parsec-corp.com/bush2
We stress the value in developing a problem statement (THIS IS WHAT OUR BUSINESS NEEDS TO FIX), then mapping data to that problem statement. If you don’t start there, you can’t really determine what it is that the data is telling you.

Once the process is begun, of course, data overload can quickly become an issue. Business leaders wonder if they’re collecting the right data. They wonder if the data is in the right place and formatted properly for use. At this point, it’s a simple data-transformation exercise to determine where the data is coming from, how it is being stored, and how insights from the data are being applied to the assets. This is often the most enlightening part of our work, as customers learn the most about themselves. This really drives to the heart of what enterprises are trying to gain from their long-term IoT / Industry 4.0 initiatives.

EMBRACING TECHNOLOGY TO ELEVATE THE HUMAN SIDE

The first genius is human, I like to say. The second genius is AI. The former excels in strategic decision-making. The latter is the hero of repeatable standards.

A harmonious partnership of intelligences is most productive; I consider the human brain a vital component to artificial intelligence. And vice versa.

Business systems deployed throughout the enterprise are more akin to islands or fiefdoms, rather than integrated operations that freely exchange critical data.
The power of real-time, data-driven manufacturing

By Greg Newman, vice president of marketing with Parsec Automation Corp.

In a world obsessed with performance, manufacturing companies must develop a solid productivity culture and strategy to successfully achieve the holy grail of continuous improvement. To meet this challenge, one of the first critical steps is to understand the roots of poor performance.

In the past, without proper technological tools, many managers relied on experience and intuition to make judgment calls. Although the value of experience must never be marginalized, considering the magnitude of investment in information technology, reliance on experience and intuition in place of reliable intelligence is risky—at best—and more than likely foolhardy. (Particularly in this era of retiring experts leaving in their wake a worrisome skills gap.)

Decision-making in the modern manufacturing world is reliant on actionable information delivered in a timely manner. The flipside of that—insufficient and untimely info—results in the flipside of continuous improvement. (Let’s call that continuous flailing.)

Imagine how many times during the course of a regular business day opportunities for making positive contributions and improving performance are squandered because decision-makers don’t have access to the right intelligence. In many cases, business systems deployed throughout the enterprise are more akin to islands or fiefdoms, rather than integrated operations that freely exchange critical data. There are many legitimate challenges in effectively integrating dissimilar systems. We all understand that. Just as the sheer magnitude of effort, expense and business disruption required to run a facility is immense.

So…what’s the answer?

A manufacturing-operations-management software solution can be the key to operating as a data-driven factory. The right software can completely change the complexion of a factory, not just from a compliance viewpoint, but also from a continuous improvement and real-time analysis viewpoint. The primary objective should be to deliver real-time, accurate, and actionable intelligence to the decision makers, while leveraging existing assets, resources and infrastructure. The reasons for this objective: quickly, economically and measurably improving operations without disruption. Or, to put it more bluntly, to make your factory make more money.

With these objectives in mind, the ideal software solution must:

• Collect data from existing systems and applications
• Allow validated and streamlined manual data entry
• Selectively and non-intrusively involve users to contextualize the collected data

Decision-making in the modern manufacturing world is reliant on actionable information delivered in a timely manner.
• Provide configurable and rules-based analytics
• Propagate the personalized business intelligence to users in real-time
• Not disrupt operations, existing applications or systems

While the ability for non-intrusive data-collection is critical, it is the seamless delivery of the information to the users (based on their personal preferences) that truly makes for a powerful decision-support tool. I recommend looking for solutions that put the exact information needed for critical decision-making in real-time, regardless of where the users are. Whether in the office, on the production floor, at home or on the road, the software should keep users informed and empowered to consistently make positive contributions to productivity.

By eliminating guesswork and reducing reliance on intuition, users will accomplish significant gains in productivity within a short time. Better decisions make a significant difference, especially when these decisions can prevent undesired events and results.

That is the holy grail. Better decisions courtesy of better data. Sound better?

*By eliminating guesswork and reducing reliance on intuition, users will accomplish significant gains in productivity within a short time.*
IIoT data is useless, but the insights are priceless

By Chris McNamara, Smart Industry content director

IIoT data is useless.

Got your attention now?

That’s the perspective of Clockwork Solutions’ Serg Posadas, VP of industry solutions, and Brad Young, director of services, who stress the value of the insights generated from data over the information itself. Take a look…

Smart Industry: Is data useless?
Serg: Data alone is not much use. To support decisions and actions, data must be analyzed and developed into informative insights that directly address a business’ objectives. And it’s much more involved than simply plugging data into a pre-fab model to create magically insightful analytics. Lots of data preparation is required ahead of the modeling. We need to understand the desired outcomes from the data analysis. Only then can the proper modeling platform, analytics, and visualization be implemented to support business goals.

We should understand that data is an overloaded term. Across an enterprise, different centers and employees define and use data with varying practices. So when we talk about data, we’re often speaking from differing bases and sometimes focused on fundamentally dissimilar goals. So as we collectively work to apply data to a specific purpose, it’s not unusual to find people pulling on different ropes. Often, data issues are complex not because the data itself is cryptic, but more because the people managing the data are complex.

This influence of the human element is changing though. Technology is evolving to enable a growing Industrial Internet of Things (IIoT) where automated data streams are gathered continuously and transmitted from machine to machine. Capital assets are being outfitted with the ability to transmit signals about their health. This acceleration of data production is fastest around the industrial data related to enterprise assets—about twice as fast as any other type of data. M2M (machine to machine) systems will communicate this data directly to each other. GE reports that within the next five years, more devices will be connected to each other than there will be people on the planet: that’s over 50 billion connected machines. This transformation will shift the complexity of data issues—from dealing with the humans in the loop to handling the soaring volume of data.

To be positioned for this future, companies must capture the value from this data—not simply explore and display it for business-monitoring purposes. An analytics platform must be in place to drive the value out of these huge data volumes by transforming historical and real-time data into data about the future and insights that directly apply to today’s difficult, complex decisions.

Brad: Data, and data points, are useless by themselves. Data combined with analytics (or data analysis) is not useless, and can be very powerful. The power is realized when smart data analysis is put into the hands of those that can make decisions to affect change. You have to have all three pieces: Data, Analytics, Decision Makers.

Smart Industry: What differentiates a piece of data from an actionable insight?
Serg: Well a piece of data is to an actionable insight what a bolt is to an aircraft engine. We need to add
lots of additional parts and processes to the bolt before it becomes part of an insight. While the bolt doesn’t make the engine, a faulty bolt can certainly break the engine. If the bolt works its way loose, the entire engine can fall apart due to the imbalance. In the same manner, analytics can be vulnerable to faulty pieces of data. It is the data scientist’s job to design models, platforms, processes and decision-support systems to withstand the problems we see with data elements every day.

Whether the data transmission went awry, or a human injected an entry error, or the data is missing, or we experience high levels of uncertainty. Our data analytics must be able to recognize these conditions and adjust to them. We don’t want to build an elegant model with elaborate dashboards only to have one problematic input drive us towards the wrong decision. The idea is to use the data to improve our business processes, not to create new problems because we failed to design our modeling and analysis properly.

One piece of data does not create an insight, so how much data is needed? Well...how much data do you have? How dirty is that data? And how much time do we have until we use that data for an important decision? You should start by clearly defining the question we’re trying to answer. Understanding the end goal comes first, before we can define the data requirements. A specific set of metrics are designed to answer each individual question. These metrics are driven by time-dependent historical observations, future states, or both. The analytics platform should then populate detailed, time-based answers to a vast set of complex questions. From these answers we construct actionable insights on business operations.

**Brad:** There is an increasingly overwhelming volume of data being collected today. There is little value in collected data unless you can use the data to obtain insights that turn into process-improving action. The bridge between data and actionable insight is analytics (or data analysis). Analytics allows you to discover patterns, trends, and correlations that describe how “things” are behaving. Leaders can take information gleaned from analytics and take action to affect the behavior of their system.

**Smart Industry:** Why is a high-level view of data critical to making it useful?

**Brad:** The whole purpose in using valuable resources to collect and analyze data is to create “actionable intelligence,” which is given to leaders so they can make informed, actionable decisions to influence the behavior of the system. As mentioned above, a high-level view of the data allows one to see patterns, trends and correlations that characterize system behavior.

**Serg:** The modeling and analysis that transforms data into information that drives decisions and actions must fit the operational situation. Are we collecting data from a process that includes lots of natural variability? Do we need to quantify and model uncertainty? Is the data being generated in large volumes? At a high velocity? Is the data format consistent or does it require lots of conditioning? Are we experiencing data gaps? Is the data accurate or is it riddled with errors? Are we dealing with many different forms of data?

By understanding the characteristics of the data available, we are able to decide which modeling platforms and analytics are suited to answer our business questions. One model cannot fit all situations, so even highly-specialized models must be customized to ingest and process the data that’s available. Also that input data will need conditioning before it can deliver analytics.

So having a good understanding of the data itself is the first step in delivering the decision support and information needed to improve business processes,
Spreadsheets weren’t designed for time series data analytics. Seeq is.

Seeq’s multiple applications enable you to rapidly investigate and share insights from data stored in enterprise data historians, such as OSIsoft PI, Honeywell PHD and GE Proficy, as well as contextual data sources, for example, SQL Server, Oracle and MySQL. Seeq’s support for time series data and its challenges – connecting, displaying, interpolating, cleansing, and contextualization – relieves you of hours and days of fruitlessly searching for insights in your process manufacturing data.

Asset Optimization
Situational Awareness
Investigation & Troubleshooting
Operational Excellence

www.seeq.com
(206) 801-9339

© 2018 Seeq Corporation. All Rights Reserved.
develop sound strategies, and adapt to changing conditions.

**Smart Industry**: Can an enterprise take full advantage without using historical data or is pairing old and new information critical to success?

**Brad**: The short answer is no—you can't take FULL advantage without using historical data. As in all cases, you have to understand your system. Part of understanding the system is knowing how it has changed over time. This allows you to provide context when pairing historical data with new data. Further, pairing historical data with new data can help you understand how your system has changed, or confirm/deny what you thought about how your system changed.

**Serg**: We can classify anything that happened in the past as historical data. So if by “new information” you mean data that was very recently collected, those inputs still fall in the historical category. Then the question becomes: How far back should we go to collect meaningful data about our assets, processes, and the operating environment?

There's no single answer. You should rely on the expertise of the data science team to determine whether data from further back in time is adding value or simply driving us further away from reality. In rapidly changing environments and seemingly chaotic operations, going further back may actually introduce errors. Yet, there is no universal rule-of-thumb. In fact the analysis team should be carefully evaluating data sources and modeling techniques before making this decision.

So if recent information, even if it was just streamed to us seconds ago, is considered historical data, then what could possibly be viewed as data that doesn't fit into this category? To capture the detailed future operations of complex systems, simulation-modeling platforms can encapsulate policies, business rules, and environmental conditions that have not been observed in the past. This type of analysis is essential if we are influencing a business strategy. The simulation generates data that did not previously exist and often generates output that is greater in volume than the data used to build the model.

Looking back at historical data enables us to react. Even sophisticated machine-learning algorithms are observing history, adjusting as new conditions are observed and offering insights based on what just happened. But no historical method can evaluate conditions and changes to operations that haven't yet been captured by historical data. So if our goal is to develop a strategy that helps us respond to new conditions and practices, we need much more than these rear-facing analytics.

For accurate results in complex, dynamic systems, we need a more complete approach that is not driven by the past. We may see outcomes that decouple the future from the present and the past. Thus, any historical model limits our view of operations to fit a rear-facing, reactive worldview is severely hampered and bound to veer off target in complex, uncertain environments.

Relying on analytics powered solely by historical data is like driving full speed ahead while staring into the rearview mirror. Your data-science team must be sophisticated enough to recognize situations that require simulation-driven results that provide insights that are influence by both the recent past and the uncertain future.

---

This technological metamorphosis has changed the relationship between field workers, manufacturers and the customers that employ them.
How tech & data are changing field work

By Indresh Satyanarayana, chief architect at ServiceMax From GE

The data revolution has transformed industries from trucking and manufacturing to retail and fashion. But one industry that has uniquely benefited from data analytics and its resulting advances? Service.

Today, connected products, mobile technology and the data that streams in from devices in the field means service workers and manufacturers can collect useful information to service machines; better anticipate and address customer needs; train new hires and grow teams faster; and power individuals to set and meet numbers-based goals.

This technological metamorphosis has changed the relationship between these field workers, their manufacturers and the customers that employ them to keep their machines up and running. As the technology continues advancing, the relationship between these groups will undergo seismic shifts as well.

RELATIONSHIPS HAVE EVOLVED ALONGSIDE DATA

With the advent of mobile technology and data-collection developed specifically for field workers, service in the field has become leagues more advanced. Field techs are now armed with the knowledge of how to approach any situation with detail, information and research on their side—the history of the machines, how they typically function and whatever outages might have occurred in the past. Essentially, data has transformed field-service workers into oracles on the go, ones who can solve problems at the touch of a button and satisfy customers on a regular basis. Mobile technology and data has taken the idea of collecting tribal knowledge and systemized it, bringing a real-time, omni-channel approach to the service-HQ relationship.

Ultimately, this has forged a first for the industry: a digital relationship between manufacturers and service workers. With real-time data flowing in, metrics like location information, schedules and availability are within arm’s reach for dispatchers at HQ, helping them be more effective in allocating resources and service in an outcomes-based model, where companies sell long-term service contracts that deliver consistent outcomes and end-results.

Imagine, Rolls Royce sells flight hours instead of engines to deliver yearly proactive service to its machines, or SolarCity operates on a power-purchase agreement to sell electricity instead of physical panels. By bringing manufacturers and service workers closer together through real-time data, this type of future is in grasp.

SMARTER MACHINES ARE HELPING MANUFACTURERS ACCOMPLISH MORE

As all of this data from connected machines accumulates and stockpiles, manufacturers can access greater insight into a machine’s lifecycle: to know in advance when it will need service. Today, machines are no longer islands—they’re fully connected. When analyzing the data that comes in, such as weather patterns causing wear-and-tear on a wind turbine, these patterns and identifying anomalies can help manufacturers better understand when an outage may occur and plan their service proactively for their customers.

This kind of analysis is known as asset-performance management—because field service is more than just attending to problems after they’ve happened, but making sure that an asset is performing at its best, constantly. But paying this close attention to a
machine's collected data lends itself to a new way of service: predictive service.

Now, whether it’s an impending issue or just a general efficiency fix, manufacturers can dispatch techs before an outage even occurs—providing a broader “health” service to customers that guarantees them a future-functioning machine for a solid amount of time. With service workers as the conduits to this new relationship between manufacturers and their customers, providing an outcome (rather than a simple one-off service) is very likely to become the new norm.

**USING THIS DATA TO SOLVE REAL-WORLD PROBLEMS**

Data is nothing without a *why*—analyzing mountains of numbers doesn’t mean anything if you’re not looking to solve a certain question. In field service, the *why* is moving toward this next generation operating model to solve larger problems.

I’ll give you an example. At a recent conference put on by my company, a field tech from medical device company Medtronic gave a presentation. He discussed how mobile technology and pre-delivered information had helped him get a radiation-therapy machine up and running in the middle of the night, ensuring that any patients lined up the next day were not impacted. At the end of the speech, he got emotional—his mother had passed away from cancer, and he wanted to make sure that everyone with the disease had a chance at remission.

Ensuring that a radiation-therapy machine doesn’t go down is truly a matter of life and death here. But imagine using technology to anticipate the (natural) outages that might occur and dispatching a tech to fix the problem between therapy sessions, so no patients are left in the lurch. Or what if Medtronics could use this technology to optimize the machines to stretch for three shifts instead of just two, before needing to recuperate? Making care and treatment more accessible is one part of this business that actually matters—and that comes from data.

Digital transformation has gotten us to this point—we’ve watched as techs have become more independent, more informed and more valued by manufacturers. We’ve also seen customers become more grateful and loyal to manufacturers and their field workers for providing real-time productivity and revenue benefits. Now, in the evolution of service as an industry, it’s time to use this digital transformation to make a real-world impact—on customers, manufacturers, and the techs that do it all. □
Digital twins for the asset operator

By Andy Bane, Element Analytics CEO, Founder and Head of Product
Sameer Kalwani and VP Engineering Sean McCormick

At its simplest, the digital twin (DT) is a dynamic digital representation of a physical piece of equipment, or thing, together with its associated environment. Every DT has a dynamic data model containing a number of data attributes of the physical thing or system it represents. Attributes are associated with sensors that measure temperature, pressure, and other variables and associated physics in order to represent real world operating conditions as well as static values like the installation date or original equipment manufacturer (OEM).

A DT can also consist of multiple nested twins that provide narrower or wider views across equipment and assets based on the process or use case. For example, a complex asset like an oil refinery can have a DT for a compressor motor, the compressor, the process train served by the compressor, and for the entire multi-train plant. Depending on its size, the refinery could have anywhere from 50,000 to 500,000 sensors taking measurements represented in the DT. At the end of the day, digital twins provide the necessary schemas required to easily compare and benchmark like things against one another—helping the user/operator to understand what’s operating well and what’s not.

THREE TYPES OF DIGITAL TWINS

As the size and complexity of DTs vary, so do functions and lifecycles. We like to think of three types:

• Status twins originate from the earliest design stages of the product cycle, mostly representing consumer products like a connected home or connected car. Data from product lifecycle management (PLM) systems is a major input, and use cases are typically device management, product control, and product quality. Most product twins have short service lives when compared to industrial assets.

• Operational twins enable industrial organizations to improve the operations of their complex plant and equipment and are used to support the work of engineers (process, reliability, etc.) and data scientists doing analytics and lifecycle operations. Operational DTs may inherit data from a status DT. Operational DTs may also include machine learning analytical models. Dan Miklovic at LNS Research calls these “Smart-Connected Asset” DTs. Operational DTs have long lifecycles, and will change over time.

• Simulation twins replicate equipment/device behavior and contain built-in physics models and even process models for what’s connected to the

The operational digital twin continues to evolve—in both complexity and capability.
equipment. Simulation-twin use cases include simulations of how equipment performs under varying conditions, training and virtual reality (VR).

Other twins (cognitive twin or autonomous twin) are beginning to receive attention, and tend to be an amalgamation of those listed above.

USE CASES FOR THE OPERATIONAL TWIN

Operations teams are looking for ways to improve asset utilization, cut operations and maintenance costs, optimize capital spend and reduce health, safety and environmental incidents. At the heart of every company’s digital transformation is a desire to achieve these objectives through analytical solutions that can augment, and even be proxies for engineers and technicians. But getting there is really hard, whether it’s working to deploy basic analytics like business intelligence, sophisticated machine-learning-driven analytics, or Industrial Internet of Things (IIoT) applications through platforms provided by OEM vendors like GE, Honeywell and ABB.

The biggest challenge is sensor data, which is mostly locked up in process historian systems and stored in a format (typically flat with no context) making analytics next to impossible. This data must be modeled and then kept continuously up to date to reflect the underlying state of affairs with the equipment and its associated asset. An operational DT solves this challenge by enabling the federation of data via the DT’s data model, which is built using metadata associated with the physical equipment.

Once this data model exists, any operations data represented by its metadata in the DT can be shipped to a data lake in an organized way for analytics. The DT’s data model can also be published at the edge to enable shipping data from one system to another, for example from a PI historian to an IIoT application like GE Predix APM. The Operational digital twin also allows for continuous maintenance of the operations data to reflect constantly changing real world conditions.

FIVE GUIDING PRINCIPLES

When it comes to building operational DTs that are both effective and sustainable, we have five ideas.

1. **Leave no data behind.** Data is the foundation for the DT, so bring it all together from the following sources:
   - Time series data from data historians, IoT hubs/gateways, and telematics systems;
   - Transactional data residing in Enterprise Asset Management, Laboratory Information Management System, Field Service Management System etc.; and
   - Static data from spreadsheets, especially those left behind by engineering firms who built the plants, and process hazard data.

   Federating as much data as possible via the DT will improve the value of IIoT analytics and applications, and also reveal new and valuable information about the physical twin that was previously unavailable. For example, a DT containing maintenance, equipment, sensor and process hazard data on a critical process can give operators brand new insights on the state of maintenance on critical equipment and how it relates to high risk process safety hazards.

2. **Standardize equipment templates across the enterprise.** The starting point for building operational DTs is the equipment template which allows for modeling the equipment, its sub-components, associated sensors, sensor attributes, and other related metadata like equipment functional location. Asset operators too often rely on the OEM model, or try to keep the model limited to only the data streams they presume they need.

   Instead, use standard templates by target equipment for every DT. This will allow for easier analytics across all compressors, pumps, motors
etc., providing a view of instrumentation coverage across each, and performance benchmarking allowing comparison of different OEMs.

3. **Insist on flexibility.** Those in industrial companies responsible for delivering operational DTs have a huge challenge. Not only must the equipment and assets be represented hierarchically (compressor motor to compressor) but also across a process. Compounding this challenge, different data off-takers want to consume the data differently, and what’s good for the reliability engineer who wants a hierarchical view of target equipment at one or multiple sites doesn’t work for the process engineer who needs to see the data across multiple units within an overall process. DTs should be flexible enough to handle both hierarchies and process representations of equipment and assets, but be able to be easily managed to meet the need of the data consumer.

4. **Integrity is integral.** Asset operators deploying analytics and applications will struggle getting their engineers to adopt these new tools if they can’t assure the integrity of the data—especially the sensor data—feeding the tools. Engineers understand the serious drift issues associated with sensor data and need to know they can trust it. Operational DTs should establish and maintain ongoing integrity of time-series data feeding twin,
including being able to manage changes to tags feeding the DT.

5. **Leverage the cloud.** Cloud-enabled data infrastructure provides the scale required to build and maintain operational DTs that can then be published to the edge for streaming analytics. The cloud also allows DTs to be built and shared across the enterprise, and enables analytics to compare processes and performance and share best practices enterprise-wide.

**FIRST-GENERATION DIGITAL TWINS**

We see DTs evolving in stages, and the good news is that through past investments many companies already have ingredients in place to begin (and many have already started and not realized) their journey.

Many asset operators already have in place status digital twins, which provide a fixed schema of how to view the current readings of a piece of equipment/device. Great examples of status digital twins in the industrial world are asset registries, control systems, process models for simulations, and multivariate process control.

All of these have been around for more than 20 years and require someone to manually create a schema of the physical environment. More often than not, operational data is stuck within a flat reference file where engineers and technicians need to memorize the name of every single instrumentation signal, resorting to sensor-naming conventions to help manage the data. Creating a status DT often required manually mapping every single sensor in the instrument registry to a process drawing, just to build a single process screen. The effort to create a new screen or report could take months of work.

Just to see “the current voltage across all pumps for the past 3 months,” one would need to build a whole new schema and data model mapping all the voltage sensor readings across tens of thousands of pumps. This would take months to build, so most assumed it wasn’t possible. Traditionally, status DTs have taken this approach wherein the use case defines the data model. This approach doesn’t scale because one can’t have a brand-new data model for every new question.

**WHAT’S POSSIBLE TODAY**

Many in the industrial sector have started down the analytics path by pursuing simulation digital twins, which require physics-based models to simulate how equipment operates. Even OEMs have begun selling these physics models as a service. When a physics model for a piece of equipment stands on its own and is not related to any other equipment, this model works incredibly well (e.g. onshore oilfield pumps, windmills, locomotives). For more complex processes and assets—a majority of industrial operations—these models don’t scale well because they’re incredibly hard to create, let alone maintain.

Accordingly, we see a different second stage in the DT evolution as operators move from basic status DTs and into a series of digital transformations of their twins. We’re often asked, “How do I extract greater value from the operational data I’ve gone to great cost to collect and store?” Our answer is to start the journey with an operational digital twin that connects a flexible, community managed data model to high-fidelity, real-time and historical data to support performing general purpose analytics in tools like PowerBI or Tableau. This avoids having to build custom data models for each analytical activity. To deliver a truly capable operational DT it must be a community-managed, flexible data model, allowing anyone (with appropriate permissions) to add their own context.

In the industrial world, the operational DT is the basis of all future work, because it is the data model. A flexible data model, along with some level of data integrity, allows operators to quickly serve data to the appropriate constituencies and applications no matter the analytical use case. All of a sudden, data becomes a massive asset for the organization and begins to take on a gravity of its own.
We’ve also begun to see a new, more advanced operational twin emerge in the market, one that includes events. An event, or labeled time-frame, unlocks machine learning (ML) capabilities. In the ML world, they call labeled data “supervised data” (ML works better with the “right data” not just “more data”). With supervised data, ML algorithms know what to look for, and can now begin to uncover insights around what may be happening during the event, just prior to the event, or even assessing what are the leading indicators of when a future event may occur.

Because the operational digital twin is the foundation of these labeled events, the ML analysis can run across thousands of sensors and over a decade’s worth of data. (Sometimes we’ve seen companies only provide events without the data model behind it—limiting analytical capabilities to just a few sensors.) With the ability to analyze thousands of sensors and decades of data, identifying what is causing an event—or predicting when a future event may occur months ahead of time—actually becomes a reality.

**TWINS IN THE FUTURE**

Once the operational digital twin is well established, the next step is to augment ML capabilities with both process and physics information—what we and others call the simulation digital twin because it simulates operations of a piece of equipment/device through physics and process models.

The best part of having an operational digital twin underlying the physics and process information is that one doesn’t need a fully comprehensive physics model. The physics and process models actually provide the necessary feature engineering required to amplify the results of ML-based analytics; the more data is layered in over time, the better the operational and simulation models become.

Based on business and operational situation, one may choose to develop a simulation digital twin before building an operational digital twin with events.

We’ve typically seen better outcomes with operational twins over simulation twins, and operational twins are easier to maintain. That said, the two together are more powerful than each in isolation.

With both operation and simulation digital twins in place, it becomes a lot easier for engineers to improve operations. The next step is to connect those improvements to the income statement and balance sheet. To improve profitability and reduce risk, we need to store the relationship between the digital twin and financial models, people, and even process hazard information.

With this information it becomes easier to use machine learning and analytics to support decisions and improve overall operations. The last stage of DT evolution that we envision is the autonomous digital twin (often called a “cognitive twin”), the overarching twin that features decision-making and control decisions managed by software. We don’t see this happening for several years, but in some circles it has come into vogue, given recent developments in the automotive sector. While a motor vehicle digital twin is much simpler than an industrial digital twin, the Society of Automotive Engineers (SAE) has published a great guide on how to think about the levels of autonomy that will emerge for industrials.

Beyond the autonomous digital twin, we may begin to see them evolve from just supporting decisions to actually making them. This digital twin must be trained in the cloud (where scalable compute is available), but be allowed to execute at the site, or “the edge,” requiring the right hardware to make these low-latency decisions.

Initially, there will still be a human in the loop, or what the SAE calls level 4 autonomy. However, as humans take the appropriate actions around ambiguous events, as more sensing is added, and as the simulation digital twins can cycle through the same scenarios hundreds of thousands of times, at some point we may begin to see full “level 5” autonomy, without humans in the loop.
Register before May 31 to save $300 off the individual rate. Even more savings when registering a group! See details at event.smartindustry.com

Learn how to formulate, implement and manage your organization’s digital strategies through:

- enabling technologies
- business process innovation
- organizational alignment

Take advantage of these key benefits and features:

- A digital strategy event solely for the manufacturing industry
- Facility tours to see technologies in action
- Unrivaled line-up of industry practitioners as speakers
- Top industrial innovators in attendance
- Myriad of networking opportunities
- Convenient location

4TH ANNUAL Smartindustry 2018 event.smartindustry.com