

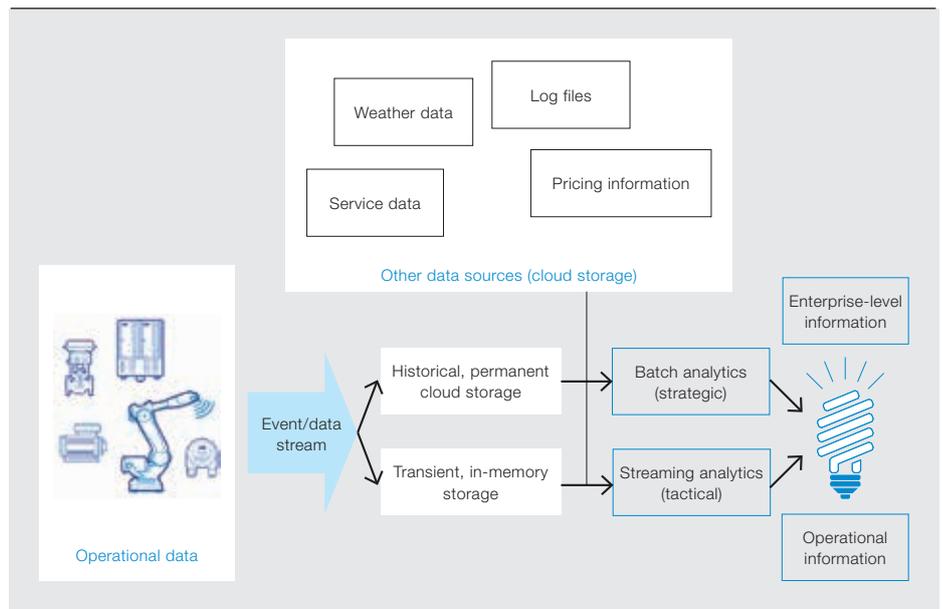


# Information exploitation

New data techniques to improve plant service

CHRISTOPHER GANZ, ROLAND WEISS, ALF ISAKSSON – Data is on everybody's agenda – and no more so than in the industrial world. Taglines such as the Internet of Things, Industry 4.0, Industrial Internet and other related topics repeatedly show up in technology publications around the world. The digitalization of the industrial world involves the collection and analysis of data from a large number of sensors with the aim of aiding the people responsible for plant operations, maintenance and management. Handling this flood of data in an intelligent way holds the key to greater plant efficiency. New techniques developed by ABB that process plant information in innovative ways can have a major positive impact on service offerings.

**1 Many sources of data in the industrial environment can be accessed to help make efficiency improvements.**



Immense quantities of data are becoming available in many industrial settings. Only when this data results in actions will efficiency gains be made.

Data flow is the lifeblood of an industrial plant. Is this the picture of the future? Partly it is, but, to a large extent, intelligent devices in plants are already communicating with each other. For example, in a typical control loop, in which sensor data is analyzed in real time by the controller and then fed back to the actuator, all the devices involved are intelligent and all of them exchange information in some form or another.

So is the future just about rearranging what already exists in a more productive way? Again, partly it is, but technologies now available allow information to be processed in new ways that can have a significant impact on service offerings.

Firstly, storing massive amounts of data has become affordable due to cloud storage offerings from several major

providers. In the industrial domain, this will allow more extensive analysis of product and system behavior than ever before. Data from multiple devices can be stored over an extended period and used to improve operations and maintenance in industrial plants – for example, by reducing asset downtime or allowing a more efficient utilization of the service workforce. Depending on the data contracts with customers, benchmark information can be shared to highlight performance deficits.

For non-real-time-critical parts of the system, this can be done now – but the quest for algorithms to uncover the most relevant insights has just begun.

Secondly, the traditional automation pyramid – with its mainly hierarchical system architecture – is being rearranged. More intelligent sensors and controllers that are more flexible exist in a meshed

network that now connects to the Industrial Internet [1]. Further, all information has to be available immediately and everywhere. Production managers want to be able to check key performance indicators (KPIs) in real time and have state-of-the-art visualization of this information on all form factors – from big-

screen, fixed installations to mobile devices like smartphones and tablets. Isolated subsystems will exist only in special cases and will suffer from reduced functionality as they will not be able to participate in the data and service ecosystem.

**Flexibility**

Flexibility is a key aspect in the new world of maximum data exploitation. One key driver of flexible systems is the need to be able to customize each and every product – rarely are two cars with the same configuration ordered during a given production cycle. This is also true for smaller, significantly cheaper products, eg, Apple Watch. Therefore, the

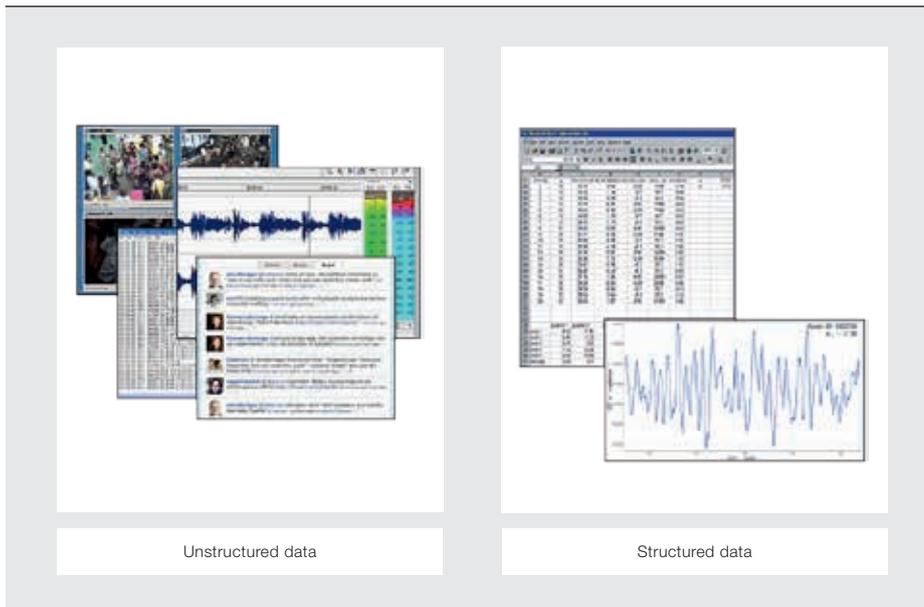
**Technologies now available allow information to be processed in new ways that can have a significant impact on service offerings.**

production facilities of the future have to be able to create products with high variability and be reconfigurable in a very short time. Further, product cycle times are getting shorter – months for smaller consumer goods and just a few years for complex goods such as cars. Such cycle times require production

**Title picture**

Immense quantities of data are becoming available in many industrial settings. Only when this data results in actions will efficiency gains be made.

## 2 The structuring of data is an important step in turning it into actions that will improve efficiency.



devices that can be plugged into the existing facility with minimal engineering. Finally, virtual engineering will need to start with early simulation models, based both on historical data and real-time data from virtual and real devices.

### Accessible data platforms

Data will no longer reside in information silos, but will be accessible for advanced analysis in cloud platforms. The analysis will access data gathered over long periods of time as well as high-frequency data streaming into the analytics engine in almost real time. A huge bonus is that these services can be scaled according to customer needs – there is no need for precautionary over-provisioning for each customer. Modern analytics platforms like Google Cloud Dataflow [2], Amazon Kinesis [3], or Spark [4] can provide the foundation for such advanced offerings.

Of course, the main role of automation suppliers is to produce the applications that are built on these platforms. These

applications can be used internally or externally. Internal applications include automated analysis of customer feedback – eg, service requests or failure reports – that improve internal processes and optimize products. External applications give customers access to advanced information – for example, operational KPIs that allow monitoring of plant productivity.

---

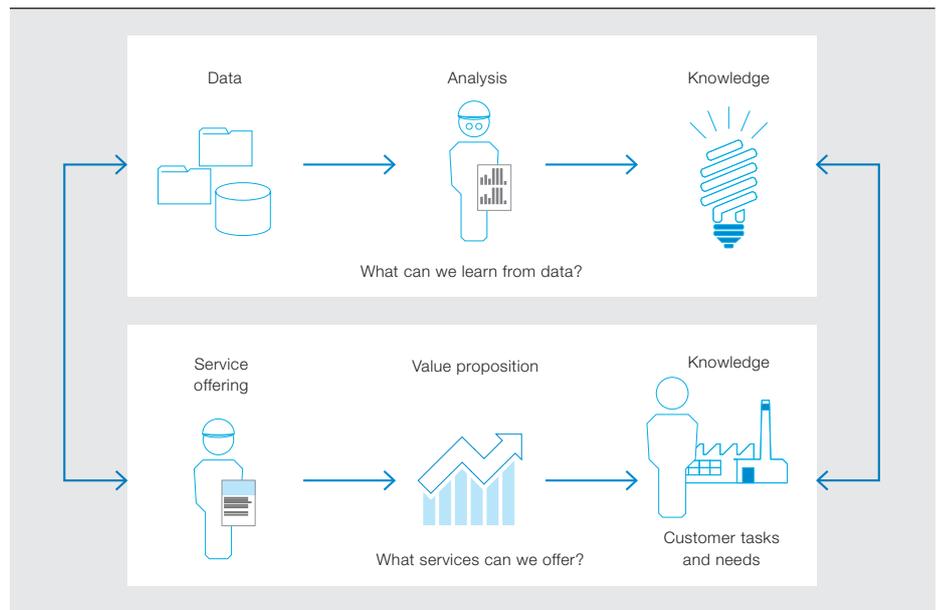
Storing massive amounts of data has become affordable due to cloud storage offerings from several major providers. In the industrial domain, this will allow more extensive analysis of product and system behavior than ever before.

### Device- and plant-level analytics versus fleet-level analytics

As already mentioned, data analysis is not new. However, many monitoring and diagnostics solutions focus on individual devices – to detect a failing sensor or to analyze vibrations from rotating machinery like motors and generators, for example. In some cases, this has been extended to entire plants or at least sections of a plant – eg, to monitor

Data will no longer reside in information silos, but will be accessible for advanced analysis in cloud platforms.

### 3 Remote, data-driven services increase plant performance and operational efficiency.



a complete shaft line with drive, motor, gearbox and load (a compressor, for instance) or to use flow sensors, pressure sensors and mass measurements to carry out leakage detection on a pipeline or water network.

The increased availability of data will enable comparisons across multiple plants – so-called fleet analytics. The fleet can be devices within one enterprise – for example, all electrical motors of a particular type. Here, a supplier like ABB potentially has access to a much larger fleet, namely the entire installed base of the motor in question. A fleet is also understood to mean the set of all complete plants of a particular type inside one corporation; all the vessels in a shipping company; or all paper machines in a paper company.

#### Homogeneous versus heterogeneous data

So far, data analysis has mainly involved signal analysis of conventional numerical process data originating from sensors. Today, there are numerous other sources of data that are waiting to be tapped. For example, search engines

are extensively used to find data on the Web, but there are many textual data sources in the industrial environment that can also be searched to yield useful results – eg, service reports, operator logs and alarm lists. Other sources include images or video files. It is usually not obvious what features to look for as the use of the data will be context- and application-dependent and this must be accommodated on a case-by-case

Like a Web search, there are many textual data sources in the industrial environment that may be searched to yield useful results – eg, service reports, operator logs and alarm lists.

basis. One challenge is to time-synchronize data from all these sources in order to fuse the information together → 1-2.

#### Edge versus cloud computing

Another interesting challenge is to identify where the data analytics should take place. The previous discussion largely assumes that the relevant data will be stored centrally, eg, in the cloud. However, devices are becoming more intelligent so there is more computational

---

# The traditional automation pyramid with its mainly hierarchical system architecture is being rearranged. More intelligent sensors and controllers that are more flexible exist in a meshed network that now connects to the Industrial Internet.

power closer to where the data is generated. To perform the computing close to the source is sometimes called edge [5] or fog [6] computing. It is already the case that not all data is sent to a data historian. For example, when using a medium-voltage drive to control the roll speed in a rolling mill, only the speed and torque are collected at the control system level, while the current is typically only available inside the drive.

With intelligent sensors and actuators, it may be that only information that has already been analyzed is available in the cloud storage. An important trade-off here, then, is to decide which signals to process locally and exactly what information to transmit to the central storage, since at the edge there is usually no historian and hence the local data may not be available for later analysis. An important factor in that consideration is that data storage costs are declining, thus reducing the need for the historian to employ compression and potentially destroy information that could be useful later.

## New insights into advanced service offerings

None of the technologies and trends described so far provides direct value to a customer. Data collection and data analysis may increase knowledge and enable predictions, but unless someone acts on these, there will be no effect on the plant performance. Only when the knowledge is turned into actions and issues are resolved will there be a benefit from analyzing more data. In other words, knowing what is faulty is one part of the equation, but fixing it is another part.

Providing remote access to data and analytics to service experts will close the loop of continued improvement. Online availability of support from a device or process expert is essential for a quick resolution of unwanted situations. Coupling remote access with the new technologies now available enables earlier detection and better diagnostics, and therefore facilitates faster service – resulting in better planning and an increase in plant and operational efficiency → 3.

### Christopher Ganz

ABB Technology Ltd.  
Zurich, Switzerland  
christopher.ganz@ch.abb.com

### Roland Weiss

ABB Corporate Research  
Ladenburg, Germany  
roland.weiss@de.abb.com

### Alf Isaksson

ABB Corporate Research  
Västerås, Sweden  
alf.isaksson@se.abb.com

---

## References

- [1] M. W. Krueger *et al.*, "A new era: ABB is working with the leading industry initiatives to help usher in a new industrial revolution," *ABB Review*, 4/2014, pp. 70–75.
- [2] E. McNulty (2014, August 8). "What Is Google Cloud Dataflow?" Available: <http://dataconomy.com/google-cloud-dataflow/>
- [3] Amazon Kinesis documentation, 2015. Available: <http://aws.amazon.com/documentation/kinesis/>
- [4] Apache Spark, 2015. Available: <https://spark.apache.org/>
- [5] H. H. Pang and K-L. Tan, "Authenticating query results in edge computing," in *Proceedings of the 20th IEEE International Conference on Data Engineering (ICDE)*, Boston, MA, 2004, pp. 560–571.
- [6] F. Bonomi *et al.*, "Fog computing and its role in the internet of things" in *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing*, Helsinki, Finland, 2012, pp. 13–16.