insight.tech

POWERING IoT AND EMBEDDED INNOVATION

# Why SCADA Is Bad for Smart Factories

A primary goal of the Industrial IoT (IIoT) is to apply enterprise analytics to operational data. But there is a problem: Industrial data is typically unstructured and timing-driven, whereas enterprise environments are highly structured and rules-driven.

This fundamental mismatch can lead to the loss of contextual and semantic information when operational technology (OT) data is transferred into IT systems.

To avoid losing this critical information, smart factories need a data architecture that bridges the divide between IT and OT environments. But analyzing the shortcomings of existing data models first will help to avoid the same mistakes.

Analytics Examples: The SCADA Setback

In traditional industrial networks, SCADA systems act as a bridge between OT and IT systems **(Figure 1)**.
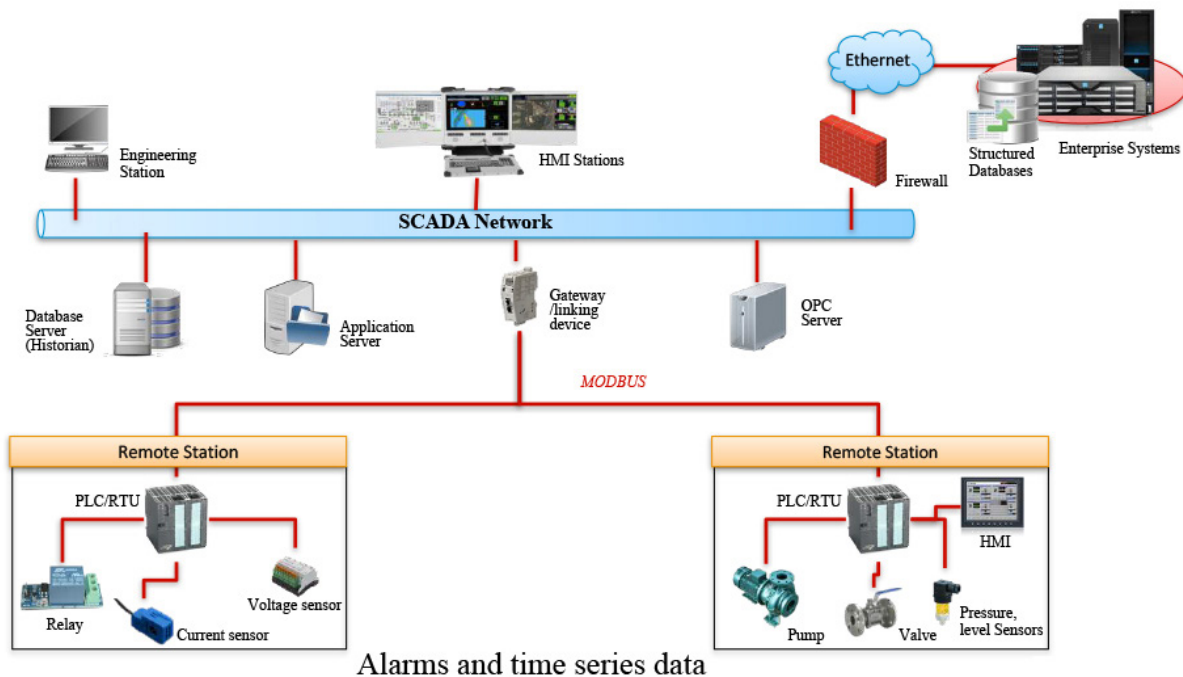
Figure 1. Traditional industrial network architectures leverage supervisory control and data acquisition (SCADA) systems to bridge OT and IT environments. (Source: Atomiton, Inc.)

Data generated by OT devices flows northbound to SCADA systems, and from there on to enterprise systems.

Because these SCADA networks are fed from the control system up, they typically rely on flow- or event-directed data architectures. In contrast, enterprise systems are based on highly structured information models governed by people and processes.

This mismatch creates a dual data architecture, with unstructured machine data on one side and highly structured systems on the other **(Figure 2)**. When OT data is sent to structured business systems, metadata such as location, time, and sequence may not be transferred.

Consider temperature data for an industrial tank. Do the reported values represent the temperature of the tank or of the fluid within? This critical piece of metadata should be retained for use in later analysis.

Not knowing the precise context around a data point can severely impact how useful the data point is. Unfortunately, attempting to retain all of this information into business process systems is incredibly complicated, both for developers and users of the system.

As a result, the value of data is reduced as it is transported across the infrastructure. Such architectures also require that analysts search for or infer the context of datasets, which further reduces the value of streaming data that becomes stale over time.

In short, SCADA systems are a poor basis for smart factories. Trying to use these networks as the backbone of an IIoT implement is essentially trying to solve a new problem with a legacy solution that wasn't effective to begin with.
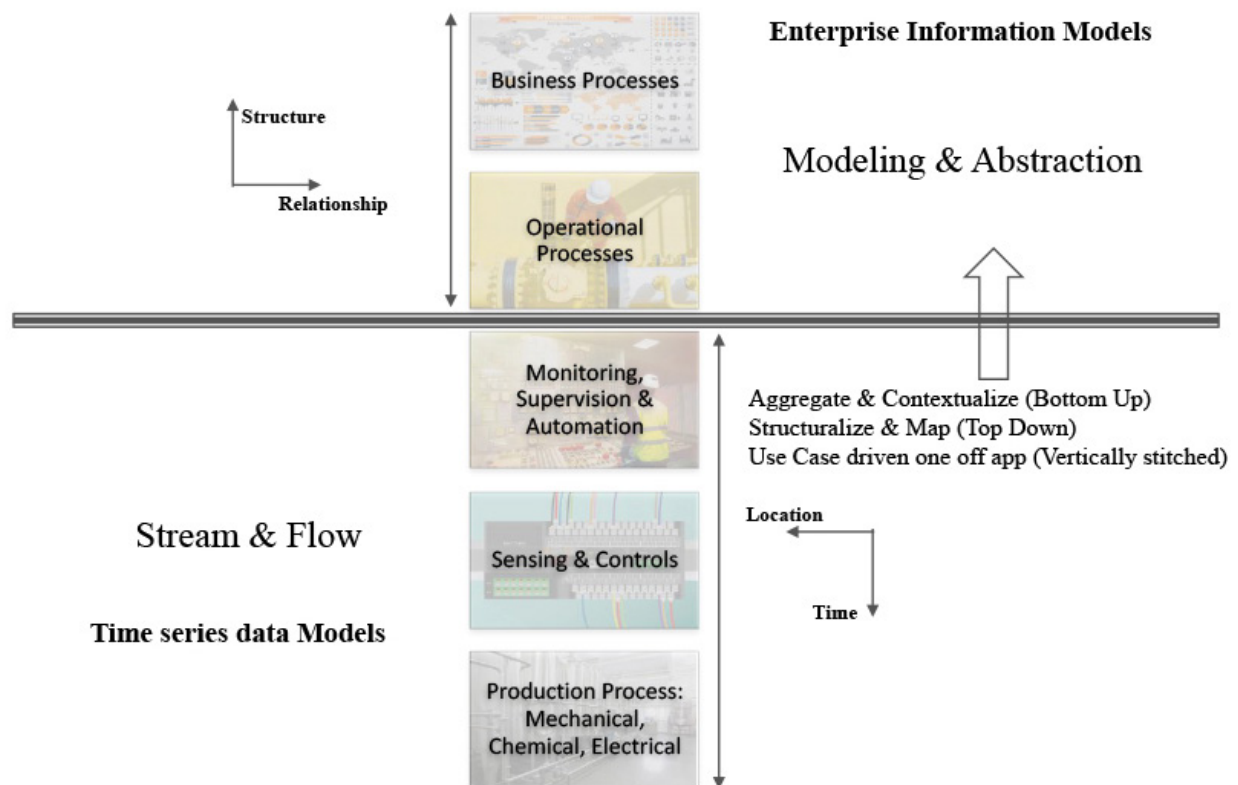
Figure 2. SCADA impose a dual-data architecture, as they can't apply enterprise information models to time-series operational technology (OT) data. (Source: Atomiton, Inc.)

Four Steps to an Effective Industrial IoT Data Analytics Strategy

Avoiding the issues of past implementations demands that infrastructure architects merge the structures and relationships of IT systems with the time- and location-centric streams and flows of industrial assets.

This can be achieved using digital models that abstract the characteristics of both data environments. Digital models are software representations of physical systems and can link the various tiers of an IIoT analytics architecture using layer-specific logic **(Figure 3)**.

Before these models can be developed and deployed, data architects need to define an analytics strategy that considers four distinct vectors:

- **Data Anatomy** involves understanding all of the data produced within a system, from end to end. The data must then be categorized based on its potential use over time for each data path.

- **Data Processing** deals with understanding the quality of data and how it can be placed in formats that enable retention of semantic value. A common taxonomy that spans all layers of the system is required to define data relationships and operational routes across infrastructure that preserve the most value.

- **Data Storage** requires knowledge of how data will be used, how often, and over what longevity, and the like. The smallest amount of data possible should be stored as close to the data source as possible, while also maintaining the data trace back to its origin.

- **Data Analytics** deal with how analytical outcomes will be used in operational and business practices. Understanding this allows the infrastructure to use available processing and storage resources as efficiently as possible so that analytics can be spread across an edge-to-cloud architecture and take place in real time on streaming data as well as offline.
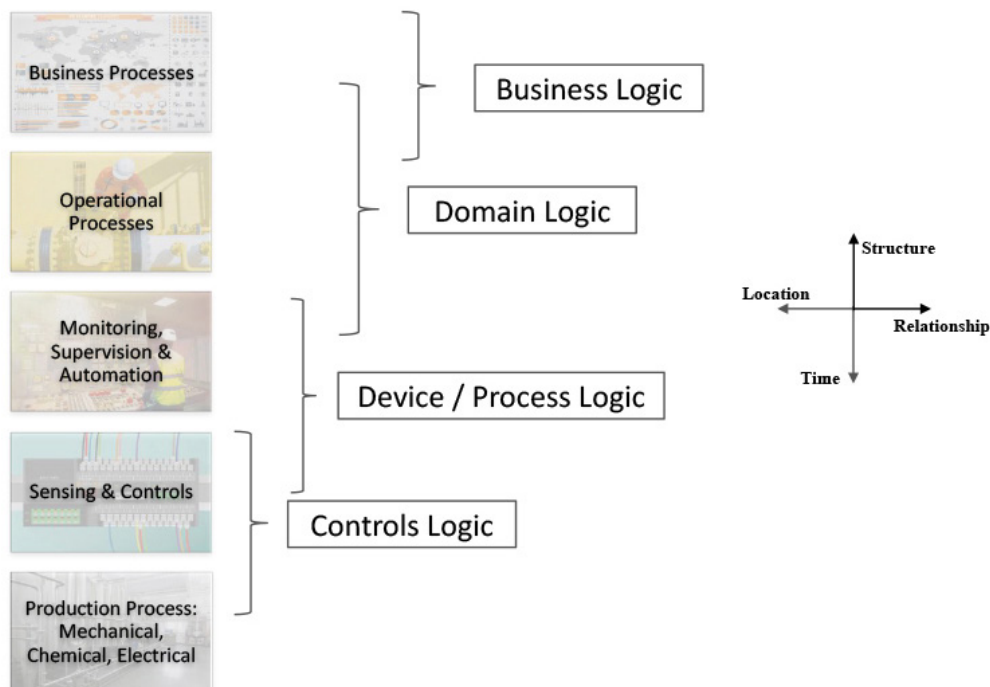
Figure 3. Digital models are software representations of physical systems that help retain valuable contextual data across the layers of an Industrial IoT data analytics architecture. (Source: Atomiton, Inc.)

Once these have been addressed, a digital model can apply layer-based logic and defined data policies that allow industrial data to be structured for enterprise systems while still retaining semantic value.

The true interoperability of data across systems also promotes higher levels of automation, as big data is contained in closed loops with the potential for ongoing analysis, learning, projections, and optimization **(Figure 4)**.
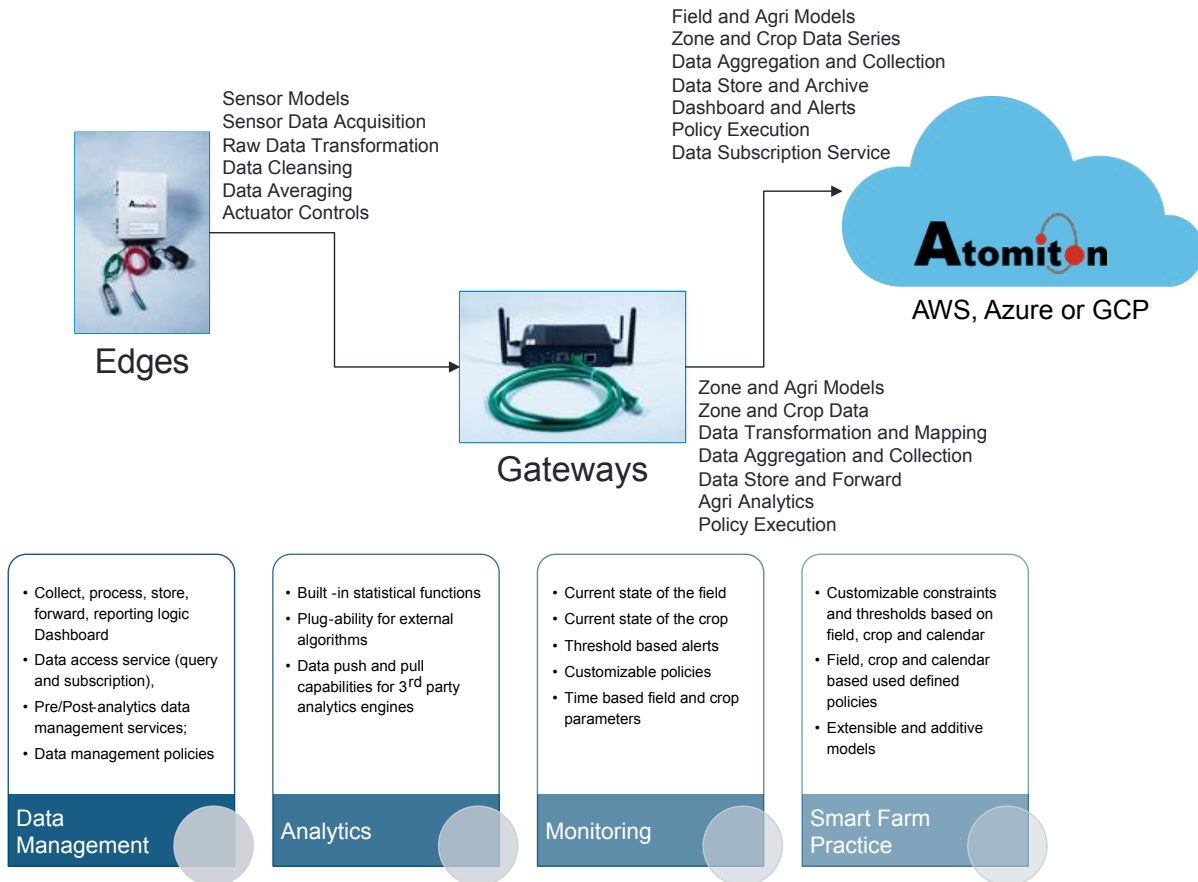
An Analytics Strategy for Smart Farming

An example of this strategy in practice is an agricultural use case jointly developed between Atomiton, Inc. and Intel®. Atomiton's TQL System is an IoT application platform comprising an IoT programming language (the Thing Query Language), a distributed container and set of services (TQLEngine), and an integrated development environment (IDE) called TQLStudio.

Each part of the Atomiton stack works to create a unified application framework that connects industrial assets and their behaviors with computing resources, while also translating data such that it can be consumed by both people and things. Intel® IoT Gateway technology occupies the role of SCADA systems in this topology, with the added benefit of 64-bit operating system support that provides smooth data transfer with cloud-based platforms.

As shown in **Figure 5**, agricultural deployments can take advantage of this architecture for improved data management, analytics, and monitoring. It sets the stage for smart farming practices that allow farmers to model their fields and crops over time and implement policies or thresholds that lead to better yields, optimized resource utilization, and the like.
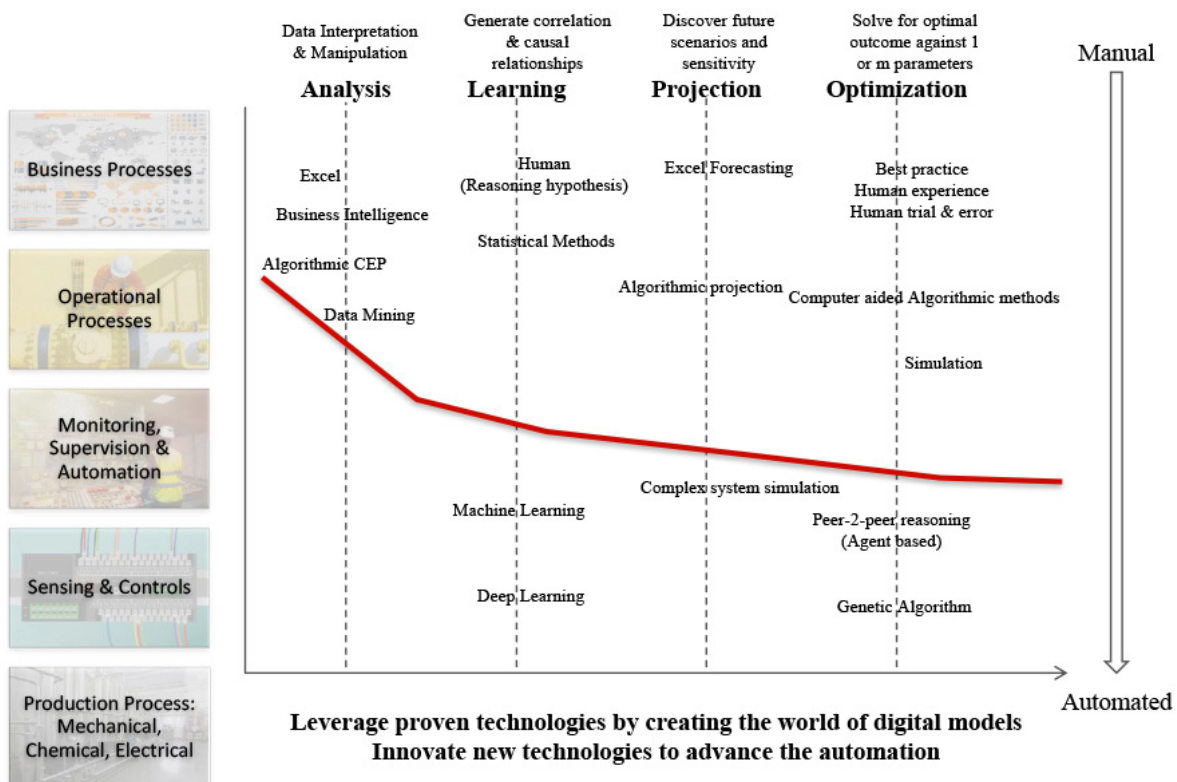
Figure 4. Digital analytics models promote a closed loop of big data that helps automate the process of analysis, learning, projection, and optimization. (Source: Atomiton, Inc.)

insight.tech

Sensor Models
Sensor Data Acquisition
Raw Data Transformation
Data Cleansing
Data Averaging
Actuator Controls

Field and Agri Models
Zone and Crop Data Series
Data Aggregation and Collection
Data Store and Archive
Dashboard and Alerts
Policy Execution
Data Subscription Service

Edges

Gateways

AWS, Azure or GCP

Zone and Agri Models
Zone and Crop Data
Data Transformation and Mapping
Data Aggregation and Collection
Data Store and Forward
Agri Analytics
Policy Execution

Data Management
- Collect, process, store, forward, reporting logic Dashboard
- Data access service (query and subscription),
- Pre/Post-analytics data management services;
- Data management policies

Analytics
- Built -in statistical functions
- Plug-ability for external algorithms
- Data push and pull capabilities for 3rd party analytics engines

Monitoring
- Current state of the field
- Current state of the crop
- Threshold based alerts
- Customizable policies
- Time based field and crop parameters

Smart Farm Practice
- Customizable constraints and thresholds based on field, crop and calendar
- Field, crop and calendar based used defined policies
- Extensible and additive models

Figure 5. Digital models based on technology such as the Atomiton TQL System and Intel® IoT Gateway Technology are enabling farmers to extract high-value IoT analytics. (Source: Atomiton, Inc.)

Crossing the Chasm of Industrial IoT Data Analytics with Digitization

While much of the above has focused on changing the way edge data is interpreted by enterprise systems, it should not be overlooked that the digitization of industrial data also has significant implications for the architecture of enterprise systems themselves.

Digital models such as those described will become more prevalent and provide better insight as to where processing and storage resources are deployed. This will drive an evolution of enterprise system architectures as well, which will not look the same even a few years from now.

With the knowledge of what it will take to cross the Industrial IoT data analytics chasm in hand, it's time to start positioning infrastructure for the leap today.