



84% of CEOs are concerned that poor data will result in lack of confidence and erode additional data initiatives.



**Data is the
new oil.**

Data Integrity: Lack of Availability and Quality

Achieving High Quality Data

Mark Brown
Yokogawa Electric Corporation
Business Development

"Data is the new oil." Given the huge expansion in analytics throughout the power industries, this expression underscores data's vastly increasing value. Unfortunately, equally common are expressions such as, "garbage in, garbage out" and, "you can have data without information, but you can't have information without data." The rampant use of such expressions underscores one simple fact: Organizations that manage data correctly and understand its inherent value have a clear advantage over organizations that don't.

High quality data not only improves decision making, compliance, and productivity, it helps maximize asset value. A recent KPMG study noted that 84% of CEOs are concerned that poor data will result in lack of confidence

and erode additional data initiatives. These executives also indicated that, in the past, poor data had affected their competitive advantage, resulted in lost revenue, and, in some cases, had caused reputational damage to their organizations.

Understanding data's value is a positive step—but how do companies manage it correctly? First, they must understand the source of the data and the path it travels to its destination on users' PC displays, smart phones and tablets. In the case of the renewable energy industry, data originates from countless sensors, data loggers, PLCs, SCADA systems, and other sources. Complicating matters is the fact that some data is invisible due to a lack of sensors or measurement instrumentation.

Other data, sourced from third party organizations, is often of questionable quality. All this makes it difficult to manage.

Therefore, it is valuable for owners to deploy standardized automation to manage data collection and reduce the overhead to staff and IT systems. However, standardized automation is often viewed as a non-essential luxury. But users who value high quality data cannot afford to pass it up. Not all assets require “rip and replace.”

Data Governance

Users who have sorted through data origins soon find out that not all data is created equal. Certain data is required at faster sampling rates than others. Some data is used to inform critical business decisions. How do companies decide on the relative value of all their data?

The development and deployment of data governance is a critical component of data integrity. This schema often contains rules for sorting data into multiple tiers ranked by importance to the organization. For example, data from a custody transfer meter at a solar facility is more important (time-sensitive) than the temperature that is measured at the back of a PLC hardware module. While both are important, one is clearly more important. The schema will contain measurement and calculation/equation standards, sampling rates, data source ID, and other key metadata that are helpful to data transparency. The template will be unique to the organization.

Another major issue is that data could be incorrect from time to time. Root causes range from instrumentation hardware failures through poorly calibrated sensors to incorrect aggregation, data conversions, or analysis. How does the user identify bad data and the actions to take on data that is suspect?

Another component that is crucial to effective data governance relates to the

Not all assets require ‘rip and replace,’ ... an overlay solution can increase data availability and reliability without the need to replace your existing hardware.

workflow for managing known bad data. There should be clear rules about validating data and how to substitute corrected values for known bad values. It should be noted that this does not include overwriting bad data. The best practice is to keep all original data, even if incorrect. Corrected data is kept on a parallel path. Any system or user accessing the data is clearly able to understand the difference between substituted values and original values as that may affect the decisions they make using that data.

Finally, it is important for data acquisition and automation technologies to be “agnostic.” In addition, it is important to ensure this technology is fault tolerant and self-aware, with the ability to recover without intervention.

Since the data that renewable assets create can be just as valuable as the kWhs they produce, companies should consider investing in a SCADA overlay solution to normalize the various technologies in their systems while increasing data availability and reliability. As asset data is used more often in downstream enterprise IT systems such as CMMS, Energy Management Systems, and Asset Performance Management solutions, tight integration with automation infrastructure will enable operators to fully realize the value of that data.

Consider investing in a SCADA overlay solution to normalize various technologies while increasing data availability and reliability.

Data governance is a critical component of data integrity