

REDEFINING PROCESS AND DATA AUTOMATION IN THE ERA OF BIG DATA



Before

e 2010, process and data automation was mostly utilized in plant or factory environments, with some innovative businesses also using it to optimize and enhance business models. But in the modern era of big data, process and data automation have really taken on a new meaning. This second post in BMC's big data series will explore the role of process and data automation in the era of big data.

Automation in IOT and Built-System Engineering

For old fashioned built-system engineers, process and data automation is nothing new. Be it in SCADA or industrial control systems, engineers have been using automation for decades to monitor, control, and report on the operations of their built-systems. But a *lack of communication standards* plagued the industry, and connections between equipment and process control systems were hard-wired, rendering process optimization efforts relatively useless. What's worse, the data produced during the course of equipment operations was meant to be used only to a limited extent, for monitoring and control purposes, and then immediately discarded.

In recent years, however, advances in information and communication technologies brought forth the *internet of things (IOT)*, where each machine in a system can be connected to, and controlled by, the internet. Similarly, big data technologies like *Apache Hadoop* have made it possible for organizations to store huge volumes of machine-generated data, thus enabling them to use that data to refine and optimize the operations of their network, and the individual pieces of equipment that comprise it. The age of real-time machine-to-machine communication and control is upon us. One impressive example of the power of IOT can be seen by looking at the tool manufacturing company, Stanley, Black & Decker. [According to Cisco](#), Stanley, Black & Decker recently connected an IOT network, thereby decreasing the complexity of manufacturing while simultaneously increasing productivity, to the tune of:

- A 10% increase in throughput
- A 16% decrease in labeling defects, and
- A 24% increase in equipment effectiveness

Process and Data Automation as a Disruptive Technology in the Business Sector

Looking now to the business sector, by the early 1990's leading businesses had begun using enterprise resource planning (ERP) software to automate business processes, optimize operations, and improve decision-support for organizational decision-makers. Unfortunately, even though ERP solutions successfully facilitated business process automation, these results often came at a *high cost*, and only *after several years* of effort. The good news is that big data technologies have helped organizations overcome these limitations.

So what's changed in the era of big data? Everything! Hadoop has made it *fast* and *affordable* for organizations to store and process petabytes of business data in a single, centralized data hub. Data ingestion, processing, and analysis can take anywhere from a few minutes to a few weeks now, depending on what big data technologies you are running. Components like Apache Oozie facilitate the automation of big data systems, thus allowing for:

- Increased operational efficiencies
- Significant decreases in big data operational costs

- Decreased reliance on human-capital, and increased IT self-reliance
- Increased scalability of big data systems due to decreases in requirements for IT resources and support staff

Also consider automated big data analytics and the benefits these have rendered, including:

- Ecommerce semantic search engines to decrease shopping cart abandonment
- Real-time fraud detection to decrease fraud incidence on ecommerce websites
- Recommendation engines to increase the number of items purchased per transaction

As enticing as all this might sound, big data automation can be complicated and there's a lot you need to know to do it right. Chapter 2 of *Managing Big Data Workflows for Dummies* will provide you the basics on what you need to know about big data workflows and processing paradigms, big data technologies, tools, vendors, and combinations thereof, as well as efficient data selection and ingestion strategies, and best practices for big data automation. To learn more about how to automate your big data systems and processes, [download the BMC Dummies guide](#) today.