

WHITE PAPER

ModelOps: Maximizing Value Across the Analytics Life Cycle

BUSINESS AND TECHNOLOGY LEADERS have bought into the concept of analytics as a lever for competitive advantage. Yet while organizations have become adept at collecting and analyzing massive data troves, many are struggling with the next step: operationalizing analytics across the entire business to capture the full value of data-driven decision-making.

For certain, there's plenty of analytics activity at the departmental level or in pilot projects across the business. However, deploying analytics efficiently and scaling efforts across the enterprise has proven to be a challenge for most organizations, primarily because of a lack of automated processes for modeling and governance. According to 2019 SAS research, organizations deploy less than half of the best analytics models they invest in developing; of those that do make it into production, an overwhelming 90% take more than three months from inception through delivery; 44% take over seven months to move into production¹.

Specific roadblocks include manual processes, inadequate data governance, and ongoing disconnects between IT and business needs. These challenges make it impossible to create a formal, end-to-end approach that supports

the entire analytics life cycle, from the earliest stages of data acquisition and discovery all the way through model development, deployment, and monitoring.

Consider the basic best practices of measuring performance of analytics models or promoting model reuse. Few organizations have taken steps to automate model performance monitoring to provide visibility into what works and what doesn't. Also rare is a universal way to recreate data features for production use, as well as a standardized way to connect model creation practices with deployment.

Organizations deploy fewer than half of their best analytics models.

Time-to-delivery for analytics models that make it into production

3+ MONTHS
90%

7+ MONTHS
44%

¹Operationalizing analytics research, SAS, 2019, https://www.sas.com/en_us/solutions/operationalizing-analytics.html

The hard truth is that analytics remains a mostly manual process at many organizations. Data silos and the burden of rework—for example, having to recreate data features or recode models in multiple languages—create inefficiencies, increase the potential for mistakes, increase costs, and ultimately limit the utility of data-driven insights.

A Path to Operationalizing Analytics: ModelOps²

Leadership teams hungry to leverage the full power of data analytics need to commit to both cultural and process change as well as new technology investments in order to optimally operationalize analytics and maximize their ROI. One modern approach that's gaining traction is "ModelOps,"

"We've reached the point in the maturity of the industry where it's not just about building models, **but about getting a return on analytics investments.**"

■ **Mark Schneider**, Advisory Product Manager at SAS.

At the same time, most analytics efforts are far too reliant on small teams of highly specialized data scientists. Tools that are too complex for the average business user limit the scalability of the model-build-deploy-monitor process, undermining the ability to "democratize" analytics to drive data-driven decision-making across the entire enterprise.

Compounding the scalability challenge is the lack of a formal process for regular model retraining. Given the rate of change in business, organizations require a systematic program to review and retrain analytics models. Without such a plan, they risk diluting the accuracy and effectiveness of insights as the models inevitably degrade over time because they aren't continuously calibrated to address changing data.

In the face of these obstacles, it's no wonder leadership teams struggle to push data analytics efforts to the next level. Most organizations are making significant investments in building powerful analytics models, but many never see the light of day due to an inability to operationalize the process.

"We've reached the point in the maturity of the industry where it's not just about building models, but about getting a return on analytics investments," notes Mark Schneider, Advisory Product Manager at SAS.

which institutionalizes best practices and technology for moving analytics models from the lab through validation, testing, and production as quickly as possible, while ensuring the highest quality results.

ModelOps takes a page from the proven DevOps playbook rooted in the software engineering world. Tailored specifically to support analytics model development, ModelOps defines a cycle of development, testing, deployment, monitoring, and retraining, allowing companies to manage and scale models to meet demand while continuously keeping tabs on them to spot and fix early signs of degradation.

ModelOps helps organizations operationalize data analytics in three primary ways:

■ Ensuring models can be deployed into production

through a number of means, including leveraging governed data, engineering for scalability, and employing tools that eliminate the translation of models into other languages prior to deployment, regardless of what language was tapped for discovery.

■ Breaking down the deployment process into

consistent, discrete steps, establishing benchmarks and metrics, and enlisting model management capabilities to automate manual tasks and reduce time to deployment.

²Learn more about ModelOps at www.sas.com/modelops.

Running the Numbers

IBM Power Systems and SAS accelerate the analytics life cycle.

2x

I/O
BANDWIDTH
vs. PCIe Gen3

1.8x

MEMORY
BANDWIDTH
vs. x86 architectures

5.6x

I/O BANDWIDTH
[CPU-GPU]
vs. compared x86 servers³

100x

EFFICIENCY
IMPROVEMENTS
for analytic modelers⁴

Instituting performance benchmarks and monitoring

to identify when a model requires retraining or retirement. Combined with integration of the training process into the model pipeline, this approach ensures maximum efficiency, fosters rapid iteration, and helps maintain quality, similar to how DevOps keeps tabs on software development.

“Analytics is at the same point that application development was 10 to 15 years ago—unable to keep up with the pace and demand to get models deployed quickly,” says Sarah Gates, Analytics Platform Strategist at SAS. “ModelOps delivers the culture, processes, and enabling technology to quickly build and deploy models, monitor performance, and then retrain, refresh or replace them as necessary to ensure firms are getting the best possible results.”

IBM Power Systems™ and SAS® Viya®: A Powerful Combination for Optimizing the Analytics Life Cycle

To support ModelOps, leadership teams will need to consider investments in hardware and software infrastructure specifically tuned to handle the analytics life cycle at scale.

Hardware: Agile and Highly Resilient

Hardware requirements include agile and highly resilient infrastructure to support mission-critical analytics. Platforms need to be flexible to adapt to the mix of computing needs associated with modern analytics, but with minimal management complexity. A purpose-built platform that supports the full analytics life cycle also must offer fast deployment options and ensure undisrupted and secure movement of massive volumes of data.

From an IT perspective, IBM Power Systems drives agility with built-in virtualization and eliminates bottlenecks with industry-leading resilience and fast data throughput. The built-in virtualization in IBM Power provides performance isolation and enhanced security. In fact, among its major competitors, PowerVM is the only hypervisor with no reported security vulnerabilities⁵.

Virtualized infrastructure reduces the cost of compute by optimizing the resources (CPU and memory) required for deployment and dynamically allocating capacity based on changing resource needs. It also helps reduce the overall infrastructure footprint by deploying multiple workloads—for example, for SAS Viya, SAS 9.4, and other data sources—in a single system. IBM Power Systems also delivers flexible models to deploy SAS Viya on-premises, in the cloud, or in a hybrid cloud.

³5.6x I/O bandwidth claim based on CUDA H2D Bandwidth Test conducted on a Xeon E5-2640 V4 +P100 vs POWER9 + V100 [12 GB/s vs 68 GB/s rated] <https://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/xeon-scalable-platform-brief.pdf?asset=14606>

⁴Learn more at www.sas.com/discover

⁵National Vulnerability Database, March 2020, <http://nvd.nist.gov/home.cfm>

IBM Power Systems and IBM Storage deliver a highly resilient environment that reduces the risk of unplanned downtime and increases the quality of service that IT can offer SAS users. Resilience and security capabilities are baked into the entire stack, helping to reduce maintenance and simplify management. The architecture has earned notable industry recognition: IBM Power Systems has been ranked first in reliability for 11 years⁶.

IBM Power Systems is also tuned for massive data throughput, accelerating data analytics by use of expansive I/O and memory bandwidth, including two times the I/O bandwidth⁷ and 1.8 times memory bandwidth⁸ versus x86 architectures. CPU-GPU interconnect capabilities usher in 5.6 times performance improvements for deep learning workloads that can take advantage of parallelization.

To complete the package, IBM serves as a single vendor for optimized compute and software-defined storage, which further helps reduce management complexity, increases scalability, and ensures the environment can easily evolve to support future needs.

Software: Open, Automated, End-to-End

Organizations also need an end-to-end, flexible analytics software platform that supports a range of users and processes throughout the information chain. The in-memory, parallelized analytics framework and support for machine learning and deep learning capabilities of SAS Viya make it easier to deliver insights to both business users and data scientists, improving their efficiency up to 100x and accelerating time to value of analytics investments.

As an open platform, SAS Viya can connect to a range of structured and unstructured data sources, including text, audio, and video, wherever it resides: in the cloud, on site, or in data lakes that are part of traditional systems. This ensures organizations can gain insights from all data—ranging from streaming to historical.

Support in SAS Viya for a range of programming languages (including SAS, Python, and R), no-code GUIs, and streamlined techniques for model building are key to breaking down silos that inhibit collaboration. The option to leverage a no-code-required visual interface makes analytics accessible to all users regardless of their analytics skills or domain experience. The ability to integrate SAS and other languages into a single, underlying code base is an insurance policy for companies worried about the attrition of critical analytics talent.

SAS Viya ensures trusted results by maintaining the appropriate governance and controls over data, models, and processes. A centrally managed inventory of analytical assets delivers lineage traceability across the entire analytics life cycle. Authentication, encryption, and authorization controls safeguard analytical assets and help enforce organizational and regulatory standards for user access.

Automated model management and deployment functionality are other tenets of operationalizing and accelerating analytics insights across the enterprise. SAS Viya delivers on this front with a range of capabilities that support model validation, monitoring of model performance, and the ability to define and execute custom workflows that drive business decisions—regardless of what language the model was written in. These are key capabilities needed for ModelOps.

Time to Act

Intelligent insights and data-driven decision-making is the lifeblood of success in the digital era. Yet without the proper infrastructure and steps to operationalize analytics efforts, the business value of data is all but lost for competitive advantage.

ModelOps provides a framework to move analytics models out of the lab and into production as quickly as possible, with quality results. IBM and SAS help organizations accelerate insights with industry-leading performance and an agile, full-stack solution that supports the full analytics life cycle with maximum resilience.

Learn more at sas.com/viyaonpower and ibm.biz/SASViya

⁶2019 Global Reliability Survey Mid-Year Update, ITIC, October 2019, <https://www.ibm.com/downloads/cas/DV0XZV6R>

⁷Pcie Gen4 provides 31.5 GB/s throughput versus PCIe Gen3 throughput of 15.75 GB/s based on publicly available material

⁸1.8X bandwidth is based on 230 GB/sec per socket for POWER9 and 128GB/sec per socket for x86 Scalable Platform