Operationalizing analytics:

What it is, why it matters and how to get started

about this e-book

The goal of analytics is to inform decision making. But for that to happen, you must operationalize it in the decisioning process. The problem is that while most organizations are able to build models, they have difficulty getting them into production. Analyst firms estimate that only 35% (IDC) to 50% (Gartner) of models are fully deployed. And SAS research discovered that 44% of models take more than seven months to deploy. It's clear that not enough models are getting deployed, and those that are take too long.

In working with our customers, we've observed many common roadblocks to successfully deploying models. These range from the lack of a data strategy to an inability to move analytics into production. In this e-book, we explore how to overcome difficulties related to data and operationalizing analytics.

table of contents

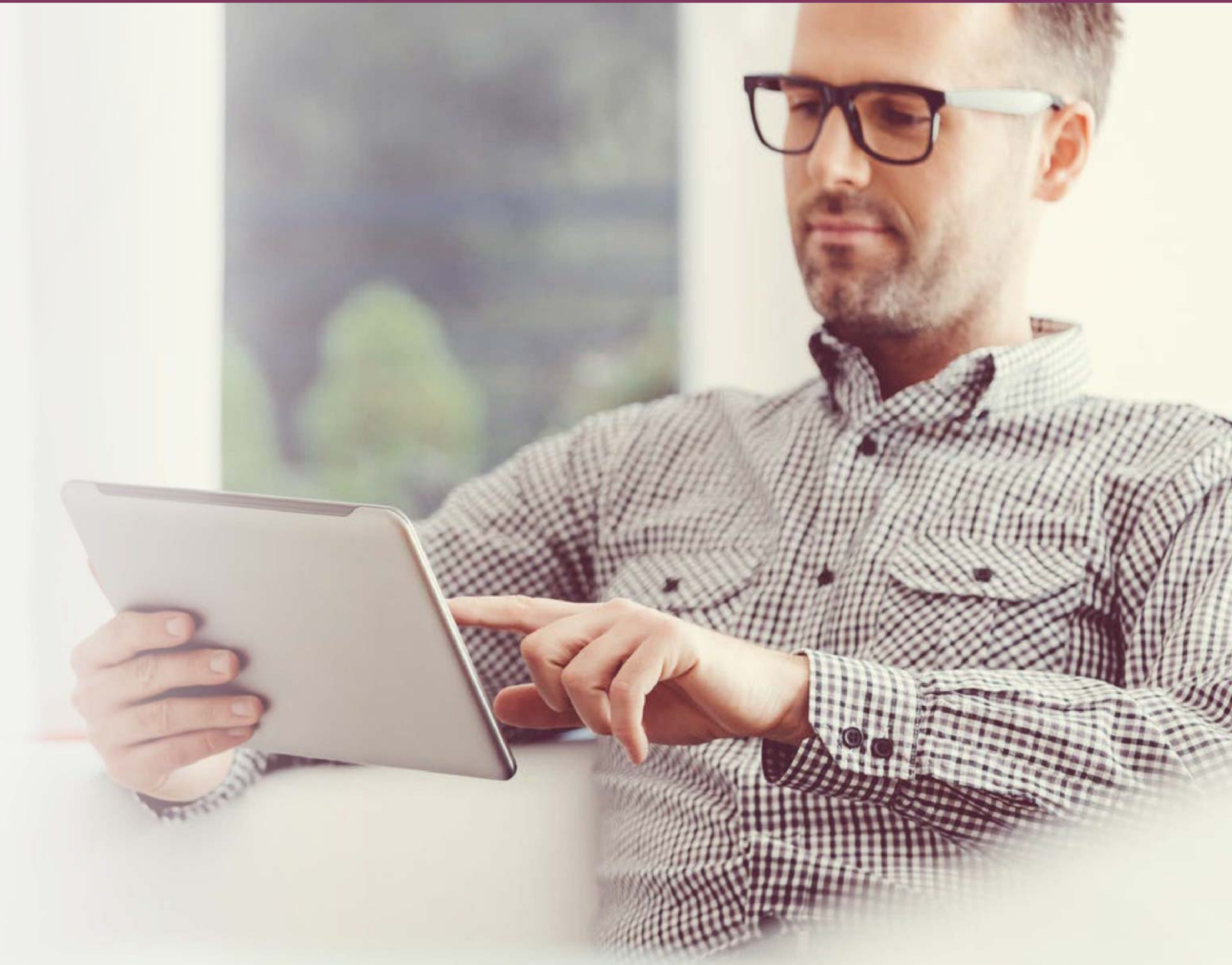
How one industry is operationalizing analytics

3 use cases from financial services

The banking industry is further along in analytics maturity than most, depending on analytical modeling to assess credit and portfolio risk, comply with government regulations, determine what financial products to offer and so much more. And yet banks are challenged to realize the full benefits of their modeling investment, especially when it comes down to the "last mile" of analytics – the gap between great analytic output and actually deriving value from it.

The No. 1 reason for missing out on their new models' potential? The lack of a repeatable process for operationalizing analytics that quickly moves models into production and monitors performance over time. According to Gartner, less than 50% of the best models get deployed, and 90% of models take more than three months to deploy. That leaves analysts frustrated, valuable data wasted and big strategic decisions in question.

The solution is a unified analytics strategy that embeds analytics in all areas of the organization to drive decision making. Here are three examples across the banking industry that show how these leading organizations followed a clearly defined path to put analytics in action to solve specific business challenges – and the results they achieved.



1

Challenge

Digital transformation

Like many banks, this Canadian financial institution wanted to transform into a digital organization – to not only build new channels to reach customers, but to also communicate with customers in their preferred channel. The problem? Disparate analytic tools and a fragmented analytic landscape.

What they did

They partnered with SAS to create a unified and agile end-to-end data science platform deployed in the cloud so that new capabilities could easily be added. By moving to a data lake, the organization minimized costly and inefficient data movement and replication.

Results

Now powered with a modernized and agile SAS® Viya® infrastructure in the Google Cloud, this financial institution can quickly model and analyze big data to improve the customer experience and produce business value.

With in-memory and in-database processing, the bank's cut processing time in half.

All users, regardless of coding language preference, can visualize analytics.

The bank has transformed into a digital enterprise that provides a top-notch, consistent customer experience across channels.

2

Challenge

Accurate risk scoring and analysis

Three regional banks in Europe were in danger of being downgraded by regulators and rating companies because they lacked a sound credit risk management process. They each struggled with lengthy time to market when changing scoring models. The analytical life cycle process for credit score models took more than a year – affecting risk in the portfolio. In addition, data gathering, analytics and reporting processes were very manual, reducing each bank's control of portfolios.

What they did

The banks implemented SAS solutions to speed time-to-market for scoring models, create a sound credit risk management process and trace changes in that process.

Results

Prebuilt data management processes helped the banks achieve a shorter time to market for credit risk models and more accurate risk scores for customers.

Improved credit risk management and flexible pricing improved the banks' ability to introduce new products quickly.

Increased accuracy and consistency of credit risk analytics – using enhanced data – benefited each bank's bottom line.

Challenge

3

Reducing operating costs and increasing efficiency

With profit margins shrinking, this global bank needed to find new revenue opportunities and reduce its operating costs by $5 billion over the next 24 months. Because of a poorly performing data and analytics infrastructure, the business was unable to respond quickly to new opportunities and regulatory requirements.

What they did

The bank used SAS solutions to migrate from its old, decentralized legacy system to new, modernized hubs, providing the bank with big data analytics capabilities as a shared service across its worldwide organization. And within the new shared services platform it implemented reporting and dashboards for multiple projects, such as international revenue, customer lifetime value, customer complaints, collections reporting and credit risk.

Results

Dramatically reduced the onboarding time of new analytics use cases.

$125 million	Saved over $125 million by consolidating the IT infrastructure to a new open source environment.
30-40%	Improved analysts’ efficiency by 30-40%.
$1 billion	Helped influence net new revenues of $1 billion.

Analytics excellence:
5 data strategy essentials

It's rare to meet a leader who isn't excited by the potential of artificial intelligence and analytics to drive digital transformation, growth strategies and operational efficiencies. But the success of an AI-based technology revolution, or just building a very simple algorithm, all begins with a solid data strategy.

A successful data strategy ensures that the organization can use, share and move data resources easily and efficiently. The goal is to guarantee that data is no longer a byproduct of business processing but a critical asset to the organization. A data strategy ensures that data is managed and used effectively and efficiently to deploy and ensure analytical models continue to perform successfully over time.

There are five core components of a data strategy that work together as building blocks to comprehensively support data management across an organization:

1 Identify

Identify data and understand its meaning, regardless of structure, origin or location. One of the most basic constructs for using and sharing data within a company is establishing a means to identify and represent the content. Whether it's structured or unstructured content, manipulating and processing data isn't feasible unless the data value has a name, a defined format and value representation.

2 Store

Store data in a structure and location that supports easy, shared access and processing. As organizations evolve and data assets grow, many have problems with the size and distributed nature of their data landscape. The goal is to store the data once and provide a way for people to find and access it. A good data strategy will ensure that all data is available for future access without requiring everyone to create their own copies.

3 Provision

Package data for reuse and distribution while providing guidelines for access. Traditionally data has been organized and stored for the convenience of the application. The problem is that most application systems are not designed to share data. The logic and rules required to decode data are rarely documented or even known outside of the application development team. In addition, most IT organizations don't provide budget or staff resources to address data sharing. Today, IT manages dozens of systems that rely on data from multiple sources to support individual business processes. Application systems and long IT processes can no longer hold data hostage. Packaging, sharing and democratizing the data is a critical shift to support the organization's success.

4 Integrate

Move and combine data residing in disparate systems to provide a unified, consistent data view. Data generated from applications is a treasure trove of knowledge, but you still have to prepare, transform or correct it before it's fit for analytics and business use. Data users need self-service tools to process data and easily deploy models while adhering to data governance and standards – without IT involvement.

5 Govern

Establish, manage and communicate information policies and mechanisms for effective data use. A governance process ensures all data constituents understand and respect the rules for shared data use. This means organizations can consistently manage data without limiting or interfering with its use. Good data governance promotes easier access, use and sharing. It also establishes trust in the data you use for the analytical and decisioning process.

Include these five components in your data strategy, and you'll have the foundation for delivering the best data for analytics and decisioning excellence. To learn more, explore SAS® data management solutions.

Why 'Ops' is in DevOps, ModelOps and DataOps

Jim Harris, Blogger-in-Chief at Obsessive-Compulsive Data Quality

What's up with Ops? It seems to be popping up everywhere these days. In fact, "Ops" is about as popular a term in the information technology disciplines now as the term "big" was over the last decade (e.g., big data management, big data quality, big data analytics, big data governance).

Think of what is probably the most well-known Ops term – DevOps. As you may know, DevOps combines software development (Dev) and IT operations (Ops) into a set of best practices intended to shorten the software and application development life cycle. DevOps uses agile methodology to focus on continuous delivery through use of on-demand IT resources and automation. And it improves the velocity, quality, predictability and scale of software engineering and application deployment. While Dev may have been the first to Ops-in (so to speak), in recent years it's becoming an Ops-tion for models and data as well (sorry for opting for continued puns).

I'm a model, but am I ready for the (production) runway?

Speaking of ModelOps: "Only about 50% of models are ever put in production," says Jeff Alford, SAS Analytics Insights Editor. "And those that are take at least three months to be ready for deployment. This time and effort equal a real operational cost and mean a slower time to value too."

ModelOps is a holistic approach for rapidly and iteratively moving models through the analytics life cycle. Whereas DevOps focuses on application development, ModelOps focuses on getting models from the lab to validation, testing and deployment as quickly as possible while ensuring quality results.

"ModelOps," Alford says, "is how analytical models are cycled from the data science team to the IT production team in a regular cadence of deployment and updates. It enables you to manage and scale models to meet demand and continuously monitor them to spot and fix early signs of degradation. ModelOps is based on long-standing DevOps principles. It's a must-have for implementing scalable predictive analytics. But let's be clear: Model development practices are not the same as software engineering best practices."

By encompassing culture, processes and technology, ModelOps can enable enterprises to efficiently and continuously develop and deploy models – not just to get more of them into production but also to deliver more advanced analytics solutions, more often.

DataOps to operationalize analytics

Borrowing methods from DevOps and ModelOps, DataOps seeks to operationalize analytics. It does so by integrating data engineering, data integration, data quality, data security and data privacy with operations to improve the cycle time of extracting value from data analytics.

DataOps facilitates "end-to-end management of ingestion, integration and utilization of data from various sources to targets," says David Loshin, President of Knowledge Integrity Inc. "Because organizations incorporate traditional (i.e., structured) data and an increasing variety of other types of information – and have to support different use cases – data integration can no longer be limited to a sequence of coordinated batch tasks of data extraction, staging, transformation and loading. Instead, organizations should introduce methods and tools to develop, manage and orchestrate data pipelines so they can develop and deploy analytics on a continuous basis."

Deploying practical artificial intelligence solutions is an enormous challenge for many enterprises. This is another area where DataOps is essential. It's "an iterative, fail-fast, learn-fast, agile process that creates proofs of value," says Kirk Borne, Principal Data Scientist at Booz Allen Hamilton. "The process involves frequent interactions between the end-users who provide requirements, the developers who build, test, and deploy the technology, and the business owners who provide metrics and evaluation."

While DataOps began as a set of best practices, it has matured to become a new and independent approach to data analytics. DataOps applies to the entire data life cycle – from data preparation to reporting – and it recognizes the interconnected nature of the data analytics team and IT operations.

Jim Harris is a recognized data quality thought leader with 25 years of enterprise data management industry experience. He is the Blogger-in-Chief at Obsessive-Compulsive Data Quality, an independent blog offering a vendor-neutral perspective on data quality and its related disciplines, including data governance, master data management and business intelligence.

3 steps for conquering the 'last mile' of analytics

Becoming insights-driven is now the ultimate prize of digital transformation, and many organizations are making significant progress toward this goal. However, putting insights into action – the "last mile" of analytics – is still a challenge for many organizations.

With continued investments in data, analytics and AI, as well as the broader availability of machine-learning tools and applications, organizations have an abundance of analytical assets. Yet the creation of analytical assets should not be the only measure of success for organizations. In reality, deploying, operationalizing, or putting analytical assets into production should be the driver for how organizations are able to get value from their AI and data science efforts.

In a traditional data and analytics continuum, data is transformed into insights to support decision making. If organizations want to break out from experimentation mode, avoid analytics assets becoming shelfware, and empower front lines to make analytics-powered decisions, they must start with decisions. Then they need to decide how to find, integrate and deliver the insights and identify data to enable that.

These days, many organizations would argue they're doing just that – they've hired analytics talent and appointed chief data officers (CDOs) or chief analytics officers (CAOs) to collaborate with business leaders to become more data- and analytics-driven. But many organizations are not seeing the desired impact and value from their data and analytics initiatives and are not able to quickly put their pilot projects into production.

According to IDC, only 35% of organizations indicate that analytical models are fully deployed in production. Difficulty in deploying and operationalizing analytics into systems or applications – and being consumed by downstream processes or people – is a key barrier to achieving business value.

Some might argue that the main focus within analytics projects has been on developing analytical recipes (data engineering, building models, merits of individual algorithms, etc.), while not much attention, priority or investment is done for operationalization of these assets. This is easier said than fixed. Data does not provide differentiation; decisions at scale do. Applying insights consistently to turn data into decisions will let organizations build a true software-led system of insights to grow and break away from competitors.

How can organizations put analytics into action in a systematic, scalable manner and conquer the last mile? Here are the three key areas where organizations need to pay consistent attention:

1

Understanding technology components

The need to streamline and operationalize model management processes requires users to register, deploy, monitor and retrain analytical models. More specifically:

Register The centralized model repository, life cycle templates and version control capabilities provide visibility into commercial and open-source analytical models, ensuring complete traceability and governance. It will also promote collaboration among different stakeholders and manage the analytics workflow effectively. Letting organizations store data, code, properties and metadata associated with models enables transparency and shows the real value of analytical assets.

Deploy The deployment step is all about integrating analytical models into a production environment and using it to make predictions. It is often the most cumbersome step for IT or DevOps teams to handle, but it's essential in delivering value. Ideally, organizations should be able to combine commercial and open source models in the same project to compare and select the champion model to deploy. Depending on the use case, models can be published to batch operational systems (e.g., in-database, in-Hadoop or Spark), on-demand systems (e.g., web applications), cloud, or a real-time system using streaming data.

Monitor Once organizations start realizing the value from analytics, the real world does not stop. Scores need to be analyzed and monitored for ongoing performance. You need to regularly evaluate whether models are behaving as they should as market conditions and business requirements change and new data is added. Performance reports can be produced for champion and challenger models using a variety of fit statistics.

Retrain If model performance degrades, organizations should take one of three approaches:

- Retrain the existing model on new data.
- Revise the model with new techniques (such as feature engineering or new data elements).
- Replace the model entirely with a better model.

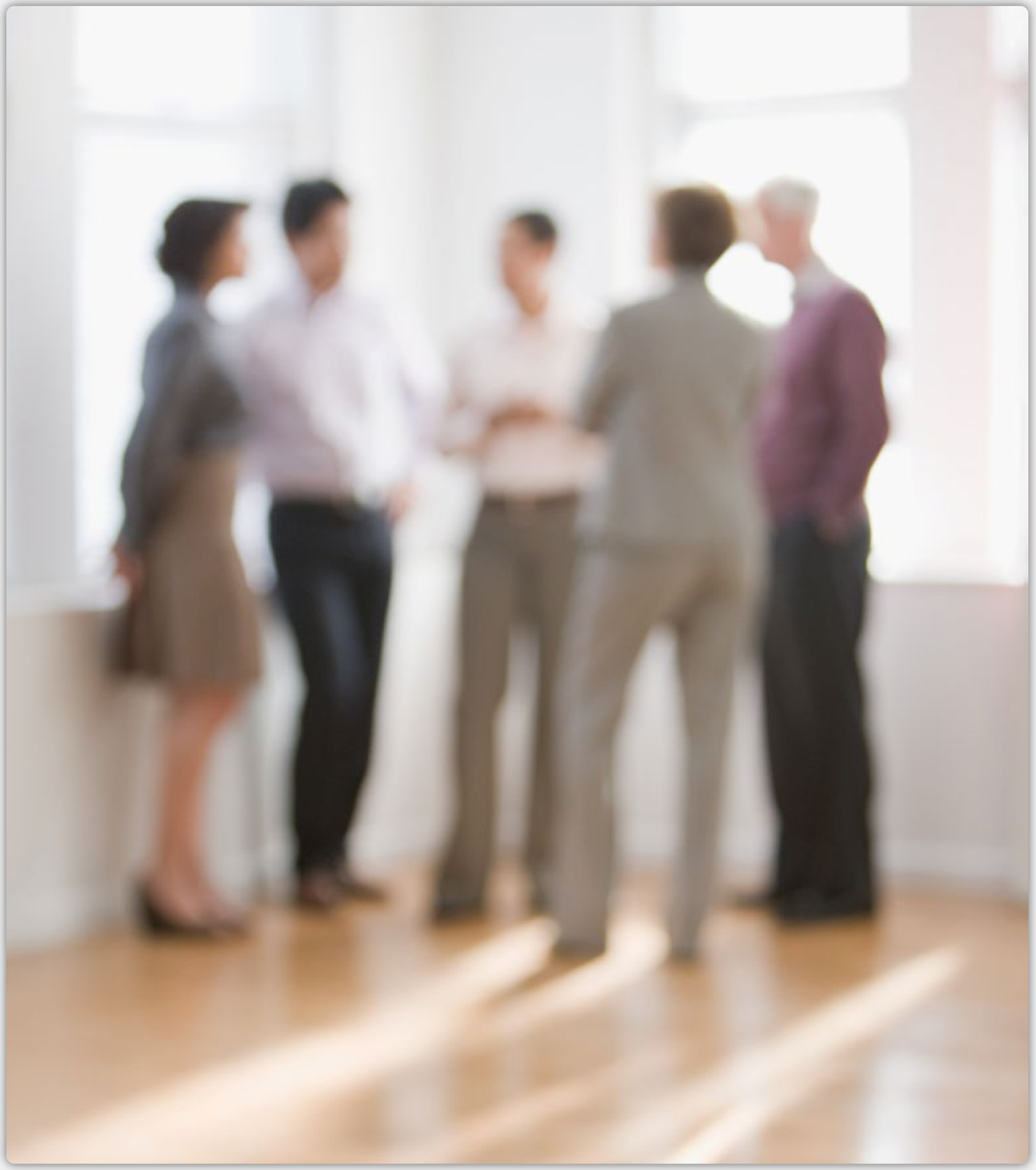
This requires commitment between stakeholders on which metrics to measure and which will deliver business impact.

2

Embracing roles and behaviors of different stakeholders

In order to be successful in the last mile of analytics, a close collaboration between stakeholders with the right skill sets – data scientists, business units, IT and DevOps – is critical. Lack of interest in dealing with deploying and managing analytics into production, leaving it solely to just one team (e.g., IT or DevOps), or not having the right incentives for all stakeholders to communicate will not create value from your analytics or AI initiatives.

For data scientists, developing analytical assets should only be initiated with deployment in mind, while IT or DevOps teams will have to understand the integration requirements, operational data flows and data preparation for model deployment and retraining. The role of business stakeholders is equally important. They are the ones who have to clearly define what benefits are expected from the analytical models and collaborate with data scientists to understand the results after models are put into production and monitor the results on a continuous basis.



3

Establishing a systematic operationalization process

Finally, the only way to ensure the value, integrity and transparency of analytical models is to establish a process for operationalizing analytics. Many organizations have a well-defined process for the analytics development phase of the analytics life cycle. But a lack of process-centric understanding around the model deployment and management phase of the life cycle is an important barrier that needs to be overcome.

A well-defined process, with proper templates and workflow, needs to validate that the model developed using training data is still performing as intended in the real world, and integrating and executing the same model against operational systems or processes. Some organizations make a mistake and stop here. In fact, to fully realize value, the performance of models in production needs to be monitored continually.

It's no surprise that this last mile of analytics – bringing models into deployment – is the hardest part of digital transformation initiatives for organizations to master, yet it's critical if they're going to experience real benefits from AI and analytics investment. To systematically realize full potential from data and analytics initiatives, organizations must involve IT and DevOps early on within the data science project such that operationalizing analytics is not an afterthought; agree on the quantifiable outcomes before building analytical models; and have a clear understanding of the steps, roles, processes and handoffs involved, from data preparation and model development to putting analytics into action.

Find how SAS can help you get from data to decisions faster with a solid data strategy and the ability to quickly operationalize analytics

Learn more

Follow us:

To contact your local SAS office, please visit: sas.com/offices