Getting started with ModelOps

A proven approach to deploying analytic models that deliver real business value

Over the last 40 years, SAS has helped clients around the world make faster, better business decisions with analytics. We've seen companies invest a great deal of time and money in developing analytics. But deployment is always a challenge – the infamous "last mile" of analytics. Lengthy delays mean wasted development efforts that never deliver their expected value.

It's time to break that cycle by redefining how we deploy models. The answer is ModelOps. It can help your organization move quality analytic models through development, validation, deployment and monitoring as quickly as possible.

Successful ModelOps requires a departure from business as usual. And change is hard. But in today's data driven, digital economy, no other endeavor delivers more value to your business than faster, more effective deployment of analytic models.

Start your ModelOps journey today with the resources in this e-book. And we'll do everything we can to help you conquer the last mile of analytics.

Shadi Shahin
Vice President of Product Strategy
SAS

TABLE OF CONTENTS

The problematic last mile:

Model deployment

Nearly every organization today is on a journey of digital transformation, which requires being able to exploit data and tirelessly apply real-time insights to drive business decisions.

The possibilities and potential business value of analytics are endless, which is why companies have spent the last decade investing in the right people, data, processes and enabling technology.

The pace of spending continues to accelerate. In 2019 alone, IDC estimates that organizations invested $189.1 billion in analytics. By 2022, they predict spending will increase to $274.3 billion[1].

However, despite such intensive investment:

- Less than 50% of the best models get deployed.
- 90% of models take more than three months to deploy.
- 44% of models take over seven months to be put into production.

<50%

Less than 50% of the best models get deployed.

90%

90% of models take more than three months to deploy.

44%

44% of models take over seven months to be put into production.

1. Worldwide Semiannual Big Data and Analytics Spending Guide, IDC. April 2019.

The lost effort and return on investment represented by the vast numbers of undeployed or delayed analytical efforts is staggering – and increasingly unacceptable to business leaders. Even when models do make it into production, handoffs and delays in moving through the analytics life cycle, combined with a lack of ongoing model management, lead to missed opportunities and decaying business value from analytics efforts (as illustrated in Figure 1 below). Clearly, getting analytical models into deployment, managing them and ensuring they continue to perform over time with new data or analytical techniques must become a top priority for organizations.

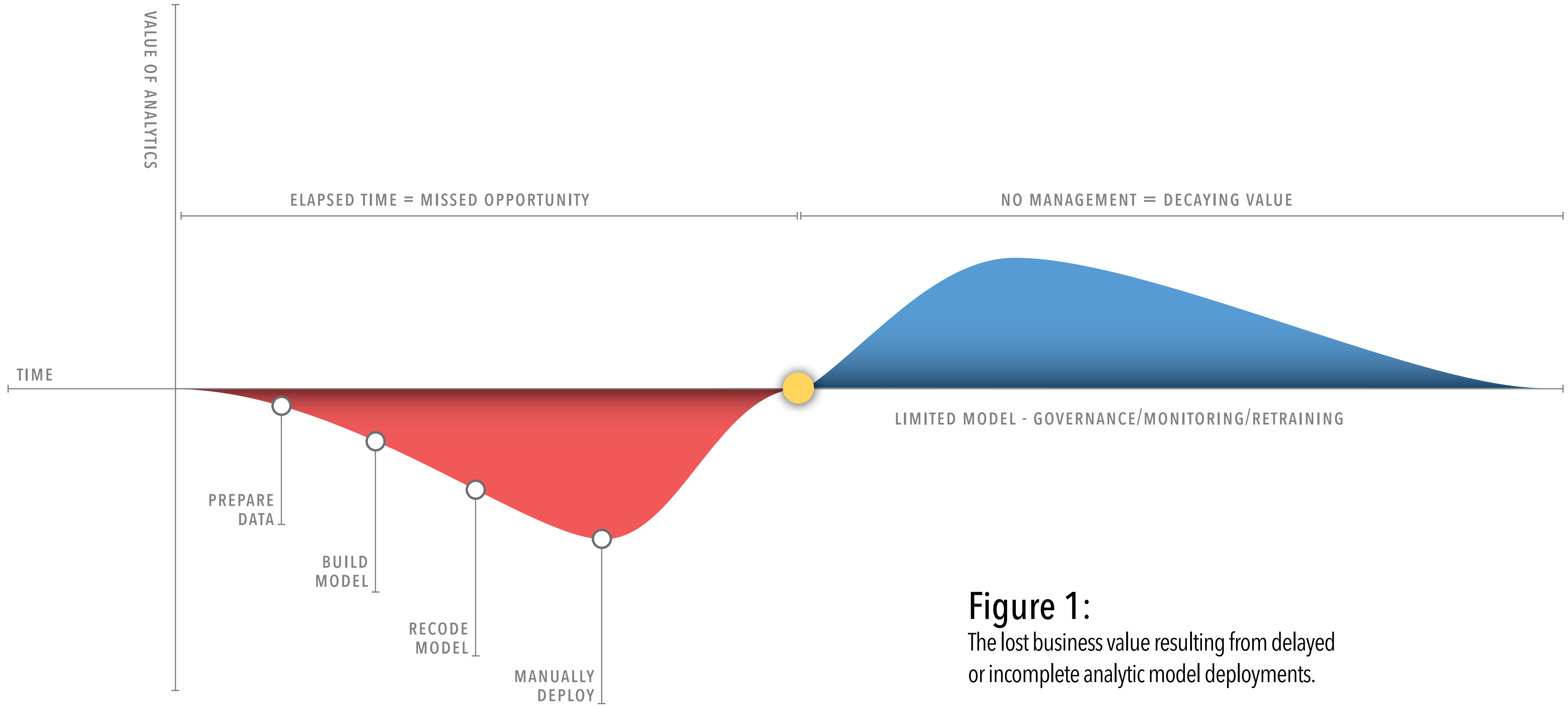


Figure 1:
The lost business value resulting from delayed or incomplete analytic model deployments.

Obstacles to model deployment

It's not surprising that model deployment is a challenge when you consider where investments have traditionally focused – for example, hiring data scientists; buying new analytic tools; and building data warehouses, marts and lakes. Unless an organization has a cohesive analytics platform or is very analytically mature, these investments in resources to operationalize analytics haven't kept up with new demands.

The recent explosion of open source tools and the ever-expanding ecosystem of technology products in use by organizations complicates deployment environments further.

Consider, for example, how data scientists tend to focus on how to build a better model or apply the latest technique. They're not typically exposed to, nor trained on, the process-centric work required to get an analytic model efficiently into production.

Likewise, IT is not typically trained on how analytic models are built and how to interpret them. As a result, when faced with complex models created by data scientists, they may have a difficult time readying them for deployment, which often leads to rework or simplification by both the data scientist and IT.

When no structured process exists for these types of activities – and there is a lack of understanding and cultural awareness for what needs to be accomplished – deployment can be very challenging.

> "The inability to put analytics into action is one of the biggest challenges across industries. To generate real business value from analytics investments, organizations need to adopt ModelOps practices."
>
> Shadi Shahin, Vice President of Product Strategy, SAS

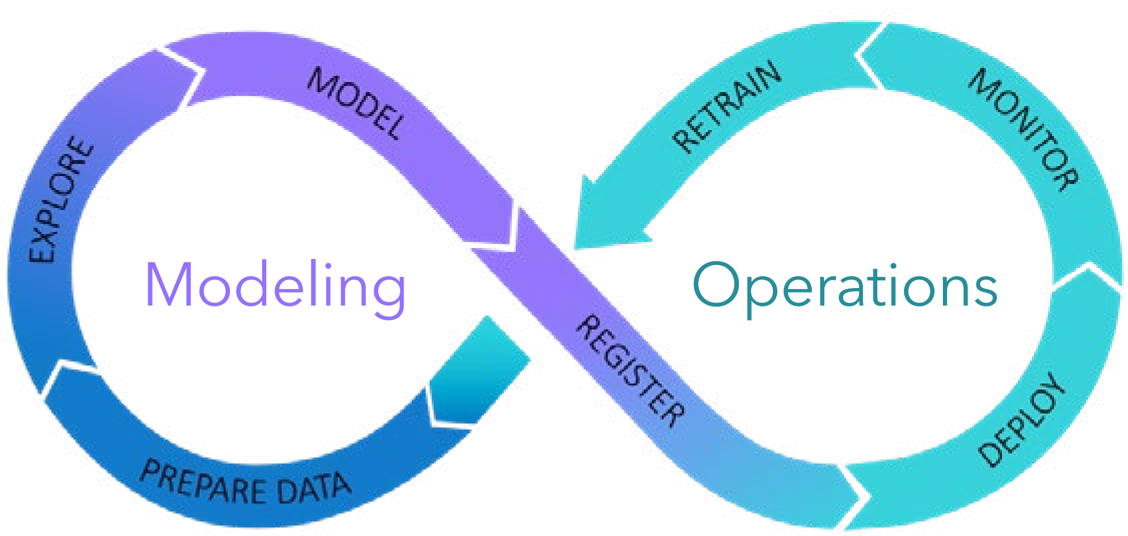
DevOps:

A model for success

Analytics teams face similar issues to those faced by application development teams over a decade ago – which were ultimately addressed by DevOps, now a mainstream way of collaborating and delivering applications faster and with higher quality to meet business needs and customer value.

DevOps emerged from problems rooted in siloed application development and operations teams, which had different and even conflicting mandates and incentives. These conflicts led to long development cycles, unhappy customers and wasted development efforts. DevOps shifted the focus of both teams from managing their respective silos to managing what needs to be delivered: an application that delivers expected business value.

When DevOps is fully adopted, all stakeholders and team members are committed to accelerating the delivery of quality software. They adopt a new cultural mindset, as well as new processes and technologies that help them deliver the right applications and services that deliver expected business value in a more collaborative, agile way. Software is developed iteratively with the customer and continuously enhanced as business and user needs change. This ensures fewer development projects fail – and when issues do occur, the team can respond swiftly to address them.



To address current challenges around operationalizing analytic models, companies need an approach similar to DevOps that brings data scientists, the business and IT operations together and empowers them to:

- Ensure the work of analytic modeling teams meets business and IT requirements so models can be easily tested and put into production.
- Vastly accelerate and scale the deployment of analytic models through different integration end points depending on the use case, such as in database, REST APIs and in stream via devices.
- Ensure models continue to perform and add value through continual monitoring and updating.

The good news is that this approach, called ModelOps, is being adopted by leading organizations.

ModelOps:

Accelerating the analytics life cycle

Similar to the way DevOps accelerates the development of applications that deliver real business value, ModelOps moves analytic models from the data science lab through registration and deployment as quickly as possible. This ensures high-quality analytic results and realization of expected business value.

At every step, ModelOps ensures that deployment-ready models are cycled from the data science team to the IT operations team in a regular cadence, and, when needed, model retraining occurs in a timely manner based on feedback received during model monitoring. In this way, ModelOps gets more analytic models into everyday use across the enterprise faster using a proven, end-to-end process that involves people and technology.

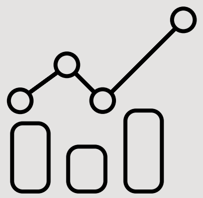
In addition, just as DevOps prevents applications from being built in a silo and "thrown over the wall" for deployment by IT, ModelOps ensures models are handed off in a manner that makes it easy for IT operations to deploy them. It also ensures that the data used to train models aligns with the operational data that will be used in production. The analytics teams and the IT organization work in harmony thanks to a mutual understanding and empathy for their counterparts and ultimate end users. This naturally influences early data and architecture choices that greatly affect the ease of future model deployment. In other words, models are:

- Developed with a deployment mindset.
- Deployed with a monitoring mindset so data scientists and analysts can monitor and retrain models as they degrade.

This approach is even more critical with the ever-expanding universe of data types, analytic techniques and languages available today. Organizations will never have a single type of model to deploy or a single source of data in use.

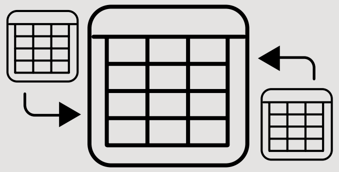
Don't forget about data

A solid data management foundation is critical to ensuring models are operationalized as effectively as possible. Some key considerations regarding data include:



Data connectivity

Connectivity and access to all relevant data that an organization needs is critical to successful model development and deployment. For models to learn and improve, for example, it takes data. Only by growing the amount of data available through more sources can success occur.



Data preparation and data quality

Data scientists are not happy or productive when they're forced to spend the vast majority of their time preparing and managing data. Organizations need data preparation tools that allow their users to quickly access data, as well as standardize, match and eventually connect to downstream applications.

Data governance

Knowing who has access to your data, where it resides, and how it's been changed over time gives data scientists the "checks and balances" they need to monitor and assess the changing quality and relevance of their data.

How is ModelOps different?

While there are many similarities between DevOps and ModelOps, it's important to understand that ModelOps is quite different due to the nature of analytical models. A traditional application is built using what is called deterministic code, meaning it does the same thing over and over; only software changes or updates can affect its behavior.

An analytical model, in contrast, is based on algorithms that predict an outcome for data that we don't currently have. These models need to be retrained, rebuilt or perhaps replaced when the relationship between the outcomes and the underlying data changes. These needs complicate the testing and deployment process and require an ongoing environment for monitoring, retraining, revalidation and redeployment. These activities make ModelOps much more complex and justified in warranting its own practice.

Realizing the benefits of ModelOps

ModelOps can transform and enhance how your business realizes value from analytics by:

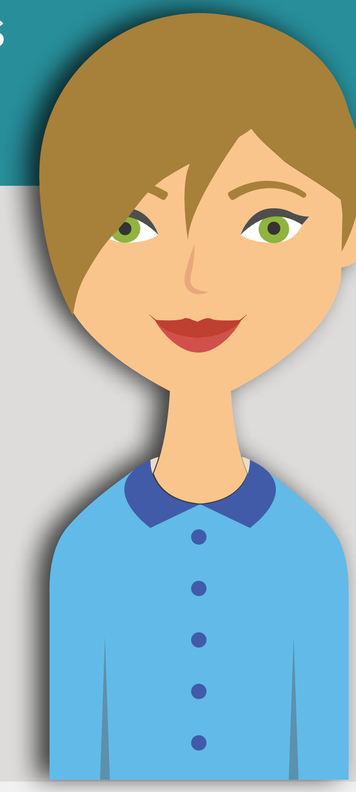
- Quickly operationalizing models to create business value sooner (see Figure 2 on page13).
- Avoiding excess complexity and recoding, which improves productivity of team members.
- Ensuring post-deployment health of models to deliver ongoing value.

By far the biggest benefit comes from being able to increase the number of models that are put into production to achieve business outcomes on consistent basis.

Equally important, this approach allows your business to invest less upfront, and yet recoup a higher value from analytics across the entire enterprise over a sustained time period. In addition, with the ModelOps approach, models continue to perform well over time, rather than decaying.

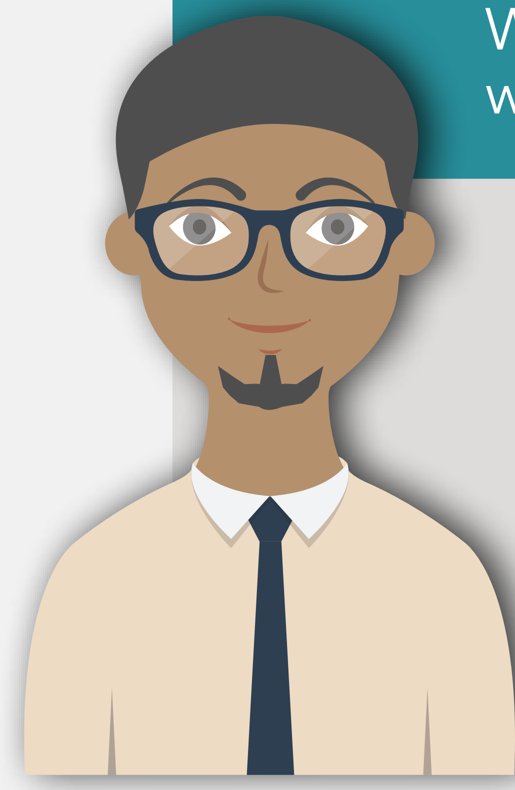
Data scientists and software developers win with ModelOps as well – in real and personal ways. Most notably, ModelOps enables both groups to work with a shared understanding of their purpose and approach, which brings clarity and context to interactions between these groups.

With ModelOps, data scientists who **build** analytic models can:



- Work in their tool of choice – open source or commercial – to solve business problems.
- See more of their models in production.
- See the full power of their models, as there's no need for IT to simplify code.
- Avoid manual, repetitive, non-value-added tasks that are inefficient and error-prone.
- Focus on delivering cutting-edge analytics for the organization.

With ModelOps, developers who **deploy** analytic models can:



- Eliminate the need for one-off, manual tasks and reinventing the wheel in order to deploy models.
- Scale model use by effectively managing and planning for capacity.
- Reduce integration challenges with well-defined models, parameters and associated data.
- Avoid becoming a bottleneck for analytics.
- Receive consistently packaged algorithm code and data.

Does my organization need ModelOps?

How organizations use analytic models changes and matures over time, but the journey is basically the same. Most start out with a few data analysts. Then, leadership starts to pay attention to the analytical results; data starts to be organized, better technology is introduced, and a few models are built.

When this investment generates business value, there's an increased demand for analytics and growing pressure for new models, faster. The challenges of data preparation, automating model pipelines and integrating champion models into production end up overwhelming analytics and IT teams.

At this point, ModelOps becomes essential. Without ModelOps, your organization's ability to efficiently move analytic models from development into deployment at scale is limited by silos, manual processes, disagreement on how value will be measured and lack of alignment by key players.

Analytics and artificial intelligence simply can't scale without ModelOps. Still wondering if you need ModelOps? Take a look at the next page to see if any of these four situations sound familiar.

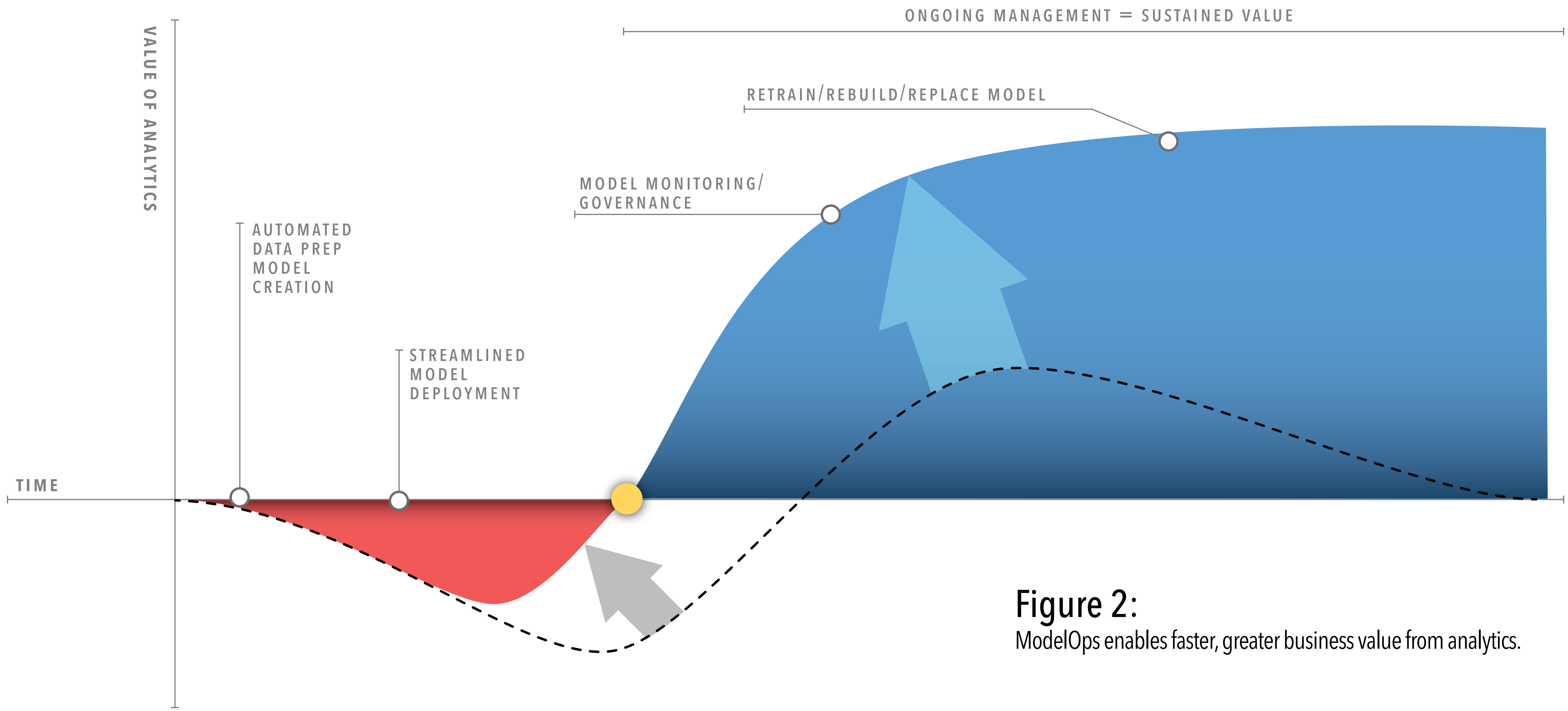


Figure 2:
ModelOps enables faster, greater business value from analytics.

4 signs you need ModelOps

BOTTLENECKS

Every day, data scientists are churning out models using open source and commercial software with a wide variety of techniques. But there are so many handoffs and points of friction between the time that the data scientist says the model is good and when it's deployed that the funnel only allows a trickle of models to get through the process. While there is the capability to build models at scale, that same capability doesn't exist in the organization to deploy, monitor and manage analytic models at scale.

BLACK BOX MODELS

You have multiple analytical models deployed in your operational systems. But no one can tell you the last time they were retrained, how well they're performing or who created them. Models are a black box that your business depends on – not a good situation. You've likely already lost the IP associated with what is deployed and will have to start from scratch in order to update the models.

TIME BARRIERS

When the business comes to the analytics team with a new problem, it takes months to prepare data, generate insights and make the model available for use. By this time, the situation or need may have changed dramatically, sending the team back to the drawing board. It feels like the analytics team is in a perpetual development cycle with nothing to show for it. The time to move models through the analytics life cycle simply can't keep up with the pace of the business.

VERSION CONTROL ISSUES

Your organization has models in production, but the processes to score data and manage versions are all manual. There are so many versions of models that it's difficult to know what version was used on what data to create what results. Data scientists frequently end up doing redundant work, and the results have to be validated in order to be trusted.

5 steps to get started with ModelOps

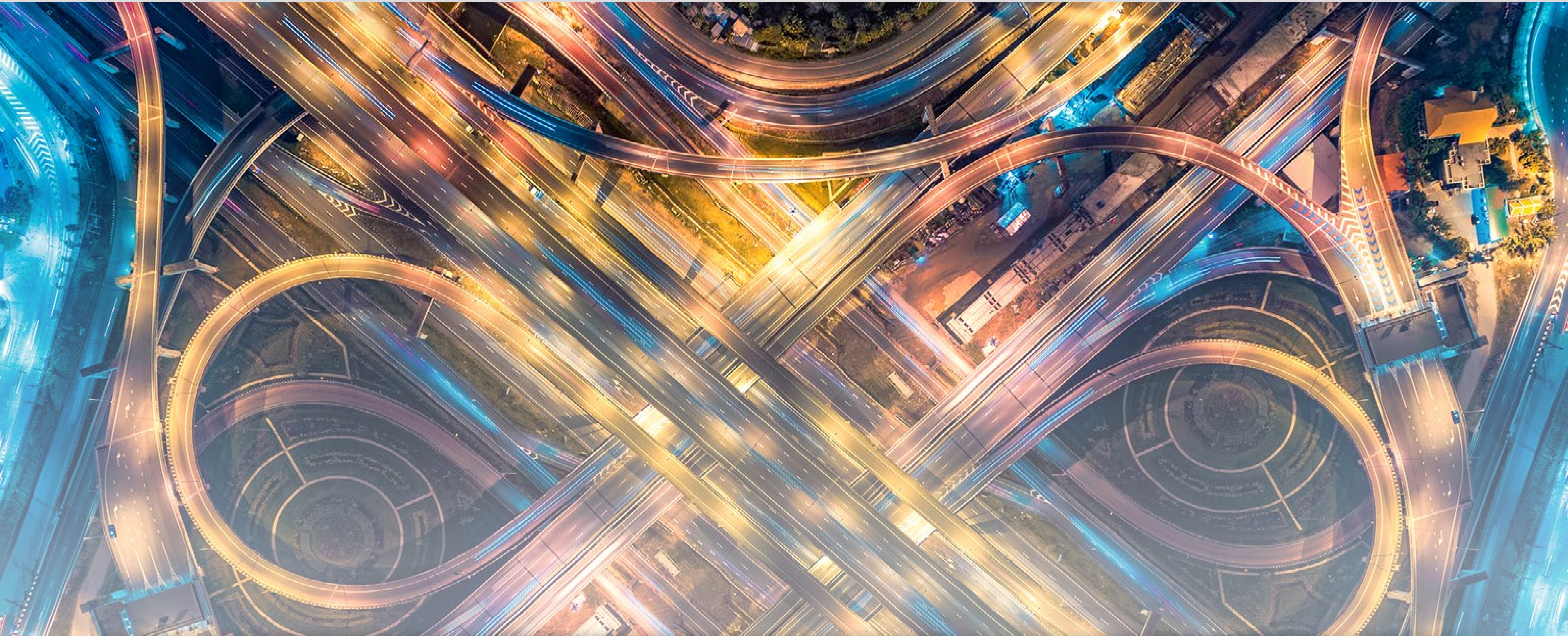
All of this sounds great, right? Clearly there's a need for ModelOps. Start by following these 5 steps:

STEP 1: **UNDERSTAND YOUR CURRENT STATE**

After you've decided to get started with ModelOps, your first step is to document your existing formal and informal analytic processes and assess how they're performing. Ask questions such as:

- How many models do we have? Where are they stored/inventoried?
- When was each model updated? By whom? How?
- Who manages our models?
- Are the right models being used in production? How do we know?
- What effort is needed to deploy models? Who's responsible? Are there documented processes?
- How long does a model take to be deployed?
- How old is the data it was trained on? Is the data clean and trustworthy?
- How are models performing?
- How do we compare different models for the same use case over time?
- Does IT work with our analytics teams to create development environments that make it possible to create models that can be easily deployed?

Asking these questions will help you identify where your bottlenecks are, which processes need to be modified or created and where to focus first. Even more importantly, these questions will trigger conversations with the people responsible for building and deploying analytics. The conversations alone will plant the early seeds of cultural change at the heart of successful ModelOps adoption.



STEP 2: **FACILITATE THE PROCESS OF TRANSFORMATION**

Once you understand the current state of your model deployment processes, focus on facilitating the cultural transformation, automation and process standardization needed to accelerate and streamline how models are built and deployed. Like most transformations, the journey will not be linear. Rather, it will be iterative and will require time and leadership for the cultural and process changes to be adopted and accepted.

It's best to start with areas of clear opportunity, make appropriate changes, demonstrate value from the changes, and continue to iterate rather than try to change everything at once. These areas of opportunity often fall into the categories of cultural commitment, standardization and automation.

STEP 3: **FOSTER CULTURAL COMMITMENT**

To move models through the analytics life cycle in a repeatable, measurable way, you need a whole new level of cross-departmental collaboration around a shared objective: creating analytic models that deliver expected business value. Everyone involved – including data scientists, data engineers, IT developers, source system experts, IT operations staff and business leaders – must collaborate around the same processes, model expectations and rules.

For this reason, you can expect that ModelOps will change how your organization operates. The silos that have created friction and handoffs in your modeling processes will be broken apart and reformed to develop cross-functional teams that deliver analytics-driven results into the hands of the business much faster. It won't be easy, but the rewards are worth the effort. As with all successful change initiatives, critical success factors include:

- Leadership from the top.
- Identifying clear areas of opportunity.
- Demonstrating value.
- Building trust and understanding.
- Aligning incentives to drive measured results.
- Empowering people.
- Adopting a mindset of iterative improvement.

STEP 4: **INVEST IN STANDARDIZATION AND AUTOMATION**

To accelerate analytic processes, you need to invest in process standardization and eventually automation – especially in the areas of implementing a model repository, performance monitoring, alerting, model deployment and retraining. With every manual task that's standardized and automated, you benefit from reduced complexity, technical debt and points of failure. As seen in Figure 3, everything from data preparation to model retirement is affected by standardization and automation.

Standardization

Standardization through ModelOps is about reducing risk and accelerating throughput by eliminating multiple approaches to the same task.

To get started on standardization, discuss and identify which processes, methods and other areas are not working well (and contributing to poor KPIs). Prioritize processes that are the most broken or will have the most impact on KPIs. Improve upon the highest-priority process, then move on to the next one.

Standardization is critical to:

- Establishing a consistent environment configuration and setup so that data scientists can drive faster model development, testing and release cycles.
- Ensuring governed data is centrally provisioned for all analytics teams, regardless of what programming language is used.
- Employing consistent, proven processes for publishing and maintaining models across the organization.
- Deploying a common repository for models used across the organization, regardless of the analytic language or technique employed.

In particular, standardization around a centralized repository is essential for enabling ModelOps.

Automation

As with process standardization, automation across the analytics life cycle can reduce risk by eliminating manual processes and enabling significant time savings – both for the analytics and the technology teams.

Key areas of focus for automation include:

- Enabling data access and preparation pipelines to streamline data provisioning.
- Automating analysis and pipeline creation for analytics and machine learning.
- Deploying and monitoring models.
- Monitoring model performance, including metrics to monitor model accuracy.
- Monitoring the change in the underlying input data feature vector distribution over time.
- Monitoring the system stability of the model.
- Triggering modeling rebuilds.
- Managing models (for example, by employing automated workflows).
- Monitoring model performance using dashboards.

It's likely that achieving these types of automation will require deploying new technology solutions or tools that are built for purpose.

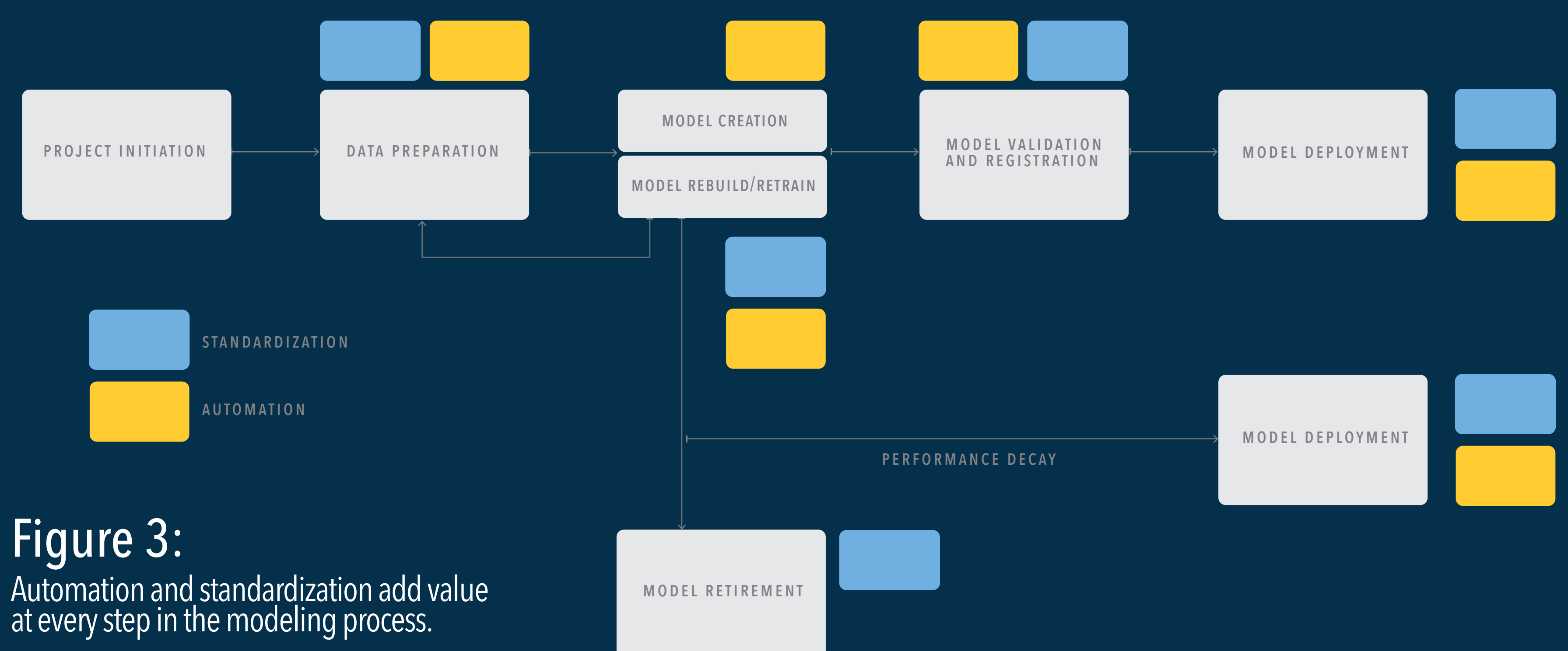


Figure 3:
Automation and standardization add value at every step in the modeling process.

STEP 5: **IMPLEMENT EFFECTIVE MODEL GOVERNANCE**

Governance is where many organizations struggle, particularly when they have multiple coding languages in use, manual processes and lack of visibility into their model deployment process. But when the three foundational requirements of ModelOps – cultural commitment, automation and standardization – are in place, you can enable truly effective model governance. For example, you can:

- **Compare models side by side** to evaluate champion and challenger models. This allows you to make comparisons between analytic models with respect to solving the best business problem, and then select the best model every time, regardless of the language used to create it.
- **Ensure the champion model is running optimally** by understanding how often it's scoring data and what version is running; automating alerts so you know when performance benchmarks are no longer met; and easily retraining, revising or retiring the model as appropriate.
- **Assess model fairness** by examining input data and/or predicted output against features that could establish or promote bias, such as race, gender, age or socioeconomic level.
- **Track the history and lineage of each model** in your organization – not just details about the model, but also what data is being used in a model and where it's being used across the business.
- **Monitor KPIs** related to the analytics life cycle to ensure processes are operating at the speed of the business and identify where they're not so they can be improved.

In other words, you gain a holistic view of all models across the enterprise, how they're performing, and what's being done to them – all without having to manually gather this data. This holistic view provides insight into the state, function, purpose, currency and value of analytic models. As a result, your organization can truly manage analytics and ensure analytic models are delivering business value.

ModelOps is a journey. Are you ready?

Like all good things, you can't implement ModelOps overnight. Adopting this approach is a journey that takes time and will present different challenges to each organization, depending on its size, existing systems and processes, and culture. But with each step, ModelOps will cultivate a new way for data scientists and IT to work together and evolve your analytics program to deliver even greater value to the business.

As you move through the steps outlined in this e-book, you'll start to evolve how teams collaborate and deploy models. Expect to see incremental improvements in both KPIs and business decision making that will have a tangible impact on the bottom line. For example, if more models are getting into production faster – and more models are being used across the business simultaneously – expect to see more, better, faster decisions based on more up-to-date models.

Similarly, by catching drift and degradation sooner through the use of new technologies and updating or repairing models faster, you'll realize more value from models in production.

Over time, continue transforming existing manual, problematic processes by deploying software that automates tasks, enables visibility, and centralizes model data and management. This transformation will drive continued improvements and accountability – and ensure everyone is focusing on the main goal: Getting more high-quality models into production quickly and making sure that they stay at peak performance to drive valuable business decisions.

Learn more about operationalizing analytics with ModelOps.

Follow us:

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