

For AI and ML Success, DataOps Acts as an Essential Accelerator

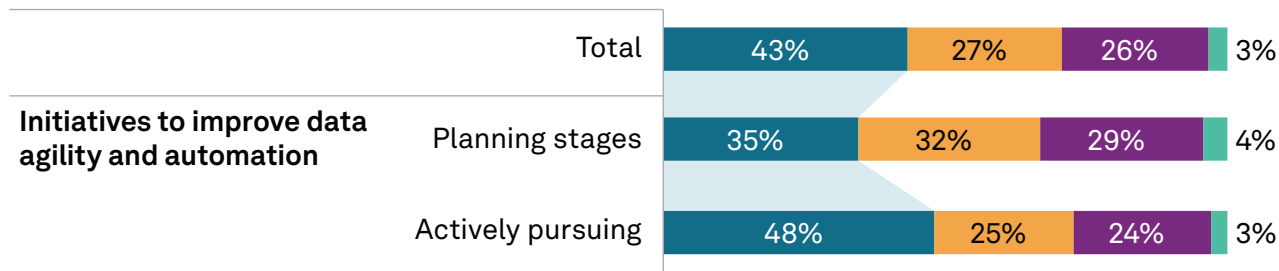
The 451 Take

Business success today, and sustained survival, relies ever more on the deft and accurate leverage of data. Both AI and machine learning (ML) outcomes depend heavily on the quality, reliability and relevance of data leveraged in training models. Yet while many organizations have a wealth of existing data, they struggle to manage that data consistently so that it is appropriate and available for use.

DataOps methodology, which is the application of more agile and automated approaches toward data management to support data-driven business outcomes, seeks specifically to provide the fundamental underpinning of reliable data flow throughout an organization. So, while analytics and data science tooling are glamorous, it is important to remember that these technologies depend heavily on a solid foundation of ongoing data management. With this in mind, organizations are focusing their efforts on DataOps as a key supporter of data-driven business outcomes — with AI and analytics outcomes emphasized.

Highest priorities for organizational adoption of DataOps

- Improving analytics/AI delivery and operationalization
- Improving data onboarding and preparation
- Improving data governance and data quality
- Measuring performance against delivery agreements



Q. What is the highest priority for the adoption of agile and automated approaches to data management (DataOps) with your organization?

Base: All respondents (n=600)

Source: Thought Leadership on DataOps, 451 Research, November 2021

Overall, the effort to improve analytics and AI delivery and operationalization is the top priority for the adoption of DataOps methodology. For organizations that are already actively pursuing initiatives to improve data agility and automation, this priority is especially pronounced. This underscores both the transformational potential of AI and analytics initiatives and their simultaneous dependence on consistent and reliable data. Given DataOps’ emphasis on the consistent availability of trusted and relevant data, it should come as little surprise that the ultimate operationalization of models is also correlated with an organization’s perceived DataOps success.

Deploying models into production is a fundamental step in reaping business value from AI and ML, yet many models never make it this far. In the same survey, participants were asked to estimate what proportion of ML models developed by their organization were ultimately deployed into production.

For organizations that reported their approach to DataOps was “falling short” of expectations, a full 20% reported that only 1-10% of models made it into production; many models were getting “stuck.” Yet for organizations that reported their approach to DataOps was “exceeding” expectations, a thin margin of 2% reported that only 1-10% of models make it into production.

Agility is also a tenet of DataOps, particularly via the use of automation in data management. This agility can translate to more rapid time to value from models. In our survey, participants were also asked how long it took — on average — for their organization to deploy an ML model into production. Again, key differences emerged based on an organization's perceived success with DataOps.

For organizations where DataOps approaches were “falling short” of expectations, only 7% responded that models took mere hours to deploy into production. Yet for organizations where DataOps approaches were “exceeding” expectations, a full 14% reported that models could be deployed to production in a matter of hours. Organizations with mature DataOps practices deploy models faster.

Business Impact

DataOps offers more agile and automated approaches to data management that provide the foundation for successful data-driven business outcomes. Initiatives such as analytics, AI and ML all depend on a supply of data that is relevant and fit for use. Proper DataOps methodology ensures a consistent “supply chain” of trusted data through an organization so that these data-driven initiatives are agile, repeatable and successful.

Regardless of DataOps maturity, the desire for analytics and AI operationalization are top motivators for DataOps initiatives. Businesses today see immense business potential with AI and analytics, yet they increasingly recognize these efforts require a foundation of consistently managed data. The highest priority for pursuing DataOps is to improve analytics and AI delivery and operationalization. This is seen not only as a business performance accelerator, but also a means to reduce operational costs.

Many ML models never deploy to production, but successful DataOps methodology can help reduce the rate of failure. In organizations where DataOps approaches are “exceeding” expectations, a much higher proportion of ML models are deployed into production where they can ultimately deliver value. The underpinning of more agile and automated data management provided by DataOps helps ensure that models are consistently trained on the right data and are fit for purpose.

Speed and agility of model deployment is correlated with an organization's success at DataOps methodology. Even if a model is deployed into production, it is not uncommon for the process to take weeks or longer. This can delay time to business value and potentially reduce the utility of the model. Organizations that report being highly successful at DataOps are much more likely to be able to deploy models into production in a matter of hours.

Looking Ahead

DataOps is not a single technology. However, technological architecture and tooling play an important role in achieving more agile and automated approaches to data management. In the DataOps journey, there is no absolute finish line, but there is a maturity curve. As organizations look to refine and improve their DataOps approaches, they need to consider how well ongoing data management practices and technology are supporting their downstream data-driven business outcomes. If data management is not actively accelerating outcomes, or somehow acts as an impediment, then practices need to be adjusted. An organization's ability to recognize its own maturity in DataOps is an important step toward improvement.

Enterprise AI and ML practices are also maturing. No longer just perceived as “science projects,” these efforts seek to deploy models directly into production where they can deliver value for the business. But with the operationalization of AI and ML, there also must be the operationalization of underpinning data management practices to deliver trusted and relevant data for models. With agility as a business focus, these data management efforts can no longer be ad hoc. Efforts need to be systematic, ongoing and iterative. A well-designed DataOps program, with appropriate supporting technology, can do this.

HITACHI
Inspire the Next

Lumada Data Ops unlocks business value by operationalizing data management with machine learning, automation and collaboration. It manages data pipelines at enterprise scale and optimizes data movements and tiering. It also simplifies data discovery and analytics to modernize data delivery processes without disrupting day-to-day business operations.

To learn more about how the Lumada DataOps platform can deliver intelligent data management across your hybrid, multicloud, core and edge business environments visit: <https://www.hitachivantara.com/en-us/products/dataops-software.html>