DataOps Lays the Foundation for Agility, Security and Transformational Change

JANUARY 2019

About this paper

A Black & White paper is a study based on primary research survey data that assesses the market dynamics of a key enterprise technology segment through the lens of the "on the ground" experience and opinions of real practitioners – what they are doing, and why they are doing it.

ABOUT THE AUTHOR

MATT ASLETT

RESEARCH VICE PRESIDENT

Matt Aslett is a Research Vice President with responsibility for 451 Research's Data, AI and Analytics Channel – including operational and analytic databases, Hadoop, grid/cache, stream processing, data integration, data governance, and data management, as well as data science and analytics, machine learning and AI. Matt's own primary area of focus currently includes distributed data management, data catalogs, business intelligence and analytics, data science management, and enterprise knowledge graphs.

Introduction

The value of data has never been more widely recognized. Enterprises of all sizes are investing in developing new data-driven applications with a view to enabling transformational change. But as enterprises strive to store and process greater volumes of data across multiple data platforms and multiple locations – both on-premises and in the cloud – so the complexities of data management and data provisioning potentially increase.

Transformational change requires greater business agility, and while application development processes have adapted to enable more rapid application development, the provisioning and management of data remains rooted in largely manual processes that can become a barrier to innovation and efficiency.

It has become clear that if enterprises are to realize business value from the development and delivery of data-driven applications and data-driven decision-making, then more agile and automated approaches to database provisioning and data management are required that are more responsive to changing business requirements. The term 'DataOps' has emerged in recent years to describe this need for more agile and automated approaches to data management.

What is DataOps?

The management and governance of data have traditionally been considered stand-alone functions, separate from (although related to) application development and delivery. This can lead to data management and governance initiatives becoming isolated from business goals, and even seen as barriers to the delivery of value.

Definition

As an umbrella term, DataOps can refer to a range of related technologies and processes, and is as much about cultural and organizational change as it is about the adoption of emerging data management products and services. For the purposes of this paper and its associated survey, DataOps is defined as follows:

> *DataOps is the alignment of people, process and technology to enable more agile and automated approaches to data management. It aims to provide easier access to data to meet the demands of various stakeholders who are part of the data supply chain (developers, data scientists, business analysts, DevOps professionals, etc.) in support of a broad range of use cases.*

DataOps is about reducing the complexity involved in data provisioning and enabling self-service access to data in order to accelerate the development of data-driven applications and data-driven decision-making, as well as supporting business agility in response to rapidly changing business requirements.

A key element of this is reducing 'data friction,' which arises when the demands of data consumers (such as data analysts, developers and senior decision-makers) are not met by data operators (e.g., data management and IT professionals). This is a perennial problem that has been exacerbated in recent years by the growing volume of data, as well as the increased number of data sources and use cases that results from escalated demand from application development and analytics projects.

The term 'DataOps' itself has not yet become mainstream. However, the associated concepts and technologies are being widely adopted by enterprises as they seek to become more data-driven. For the purposes of the survey associated with this paper, we asked respondents to consider any technologies or initiatives that seek to improve data management as DataOps (whether or not they are referred to as such within the respondent's organization).

Impact of DataOps

As stated, adoption of the term DataOps is not particularly widespread. However, 451 Research has seen growing adoption of the underlying concepts. The results of our DataOps Survey indicate that the enterprises that have embraced those associated concepts are enjoying the benefits of doing so, and are investing in further adoption of DataOps as a result.

- Almost all respondents – 92% – agree to some extent with the statement that improved DataOps would have a positive impact on their organization's success.
- 89% of respondents expect to significantly or moderately change their processes, organizational structure or training/consulting to improve DataOps in the next 12 months.
- 86% of respondents plan to increase spending, investment or development effort on DataOps technologies in the next 12 months.

Figure 1: The Positive Impact of DataOps
Source: 451 Research DataOps Survey
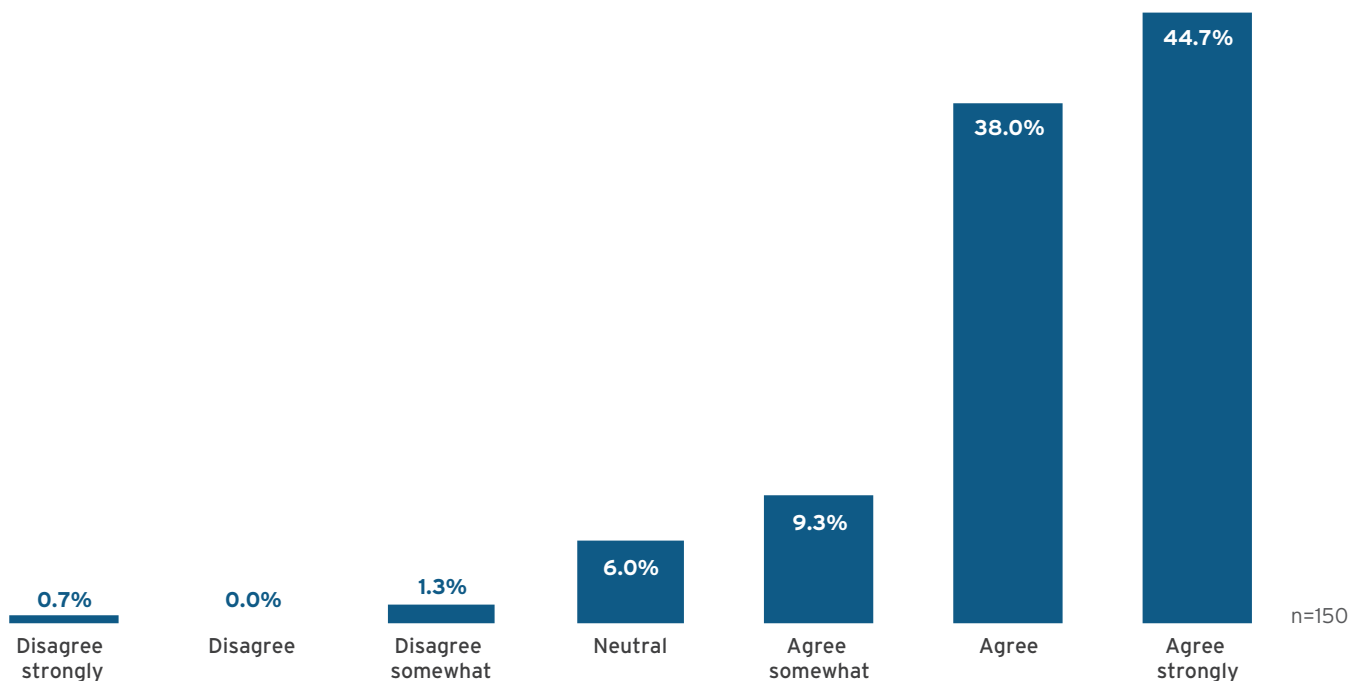
The value of DataOps has been most visible in DevOps environments where velocity is paramount and developers have a growing role in determining data access and usage requirements, as well as influencing the choice of more agile data management products and services. However, we see that DataOps is also penetrating other business processes, driven by the similar growing influence of data scientists and data engineers, for example.

As such, DataOps is also related to the need for new cultural and organizational approaches to data management associated with the shift toward being more data-driven. This includes the formation of cross-functional collaborative teams that combine data scientists, data engineers and data analysts along with data management and IT professionals.

The DataOps Survey results indicate that these cultural and organizational approaches are being adopted across a range of industries including technology, banking and financial services, and manufacturing, and also that respondents perceive that there are a range of business activities that could benefit from DataOps.

These include cloud migration (cited by 65% of respondents), business intelligence/analytics (63%), and backup and recovery (63%), as well as compliance and risk management (59%), development and testing (59%), and also artificial intelligence and machine learning (63%). However, security and compliance are in fact the biggest driver, the most significant priority, and the highest perceived benefit of DataOps (see section 'How DataOps Can Help').

BLACK & WHITE | DATAOPS LAYS THE FOUNDATION FOR AGILITY, SECURITY AND TRANSFORMATIONAL CHANGE

The survey results also indicate a close correlation between investment in DataOps and digital transformation projects, to the extent that enterprises that have invested in DataOps are further ahead in the digital transformation journey and are better placed to outlive those that have not, in the Darwinian survival of the fittest that is the drive to digitally transform business processes and organizational activities.

Among those organizations that are 'Optimized' for DataOps (with DataOps ingrained in company culture – see Demographics & Figure 7), 94% are already executing digital transformation strategies, compared with 67% of all respondents. This is perhaps not surprising given that there are common traits between digital transformation and DataOps, such as the alignment of IT with business strategy, and the continuous improvement in business operations.

Why is DataOps important?

451 Research believes that an agile approach to data management is critical if enterprises are to realize the benefits of investing in data processing and analytics technologies to develop new applications that deliver improvements in operational efficiencies and competitive advantage to drive transformational change. Being data-driven is a goal for most enterprises, but is more difficult to achieve in practice given the need to balance the requirements of various users (data analysts, data scientists, developers, IT operations and business executives for example) and sometimes competing requirements (such as agility, security, privacy and compliance).
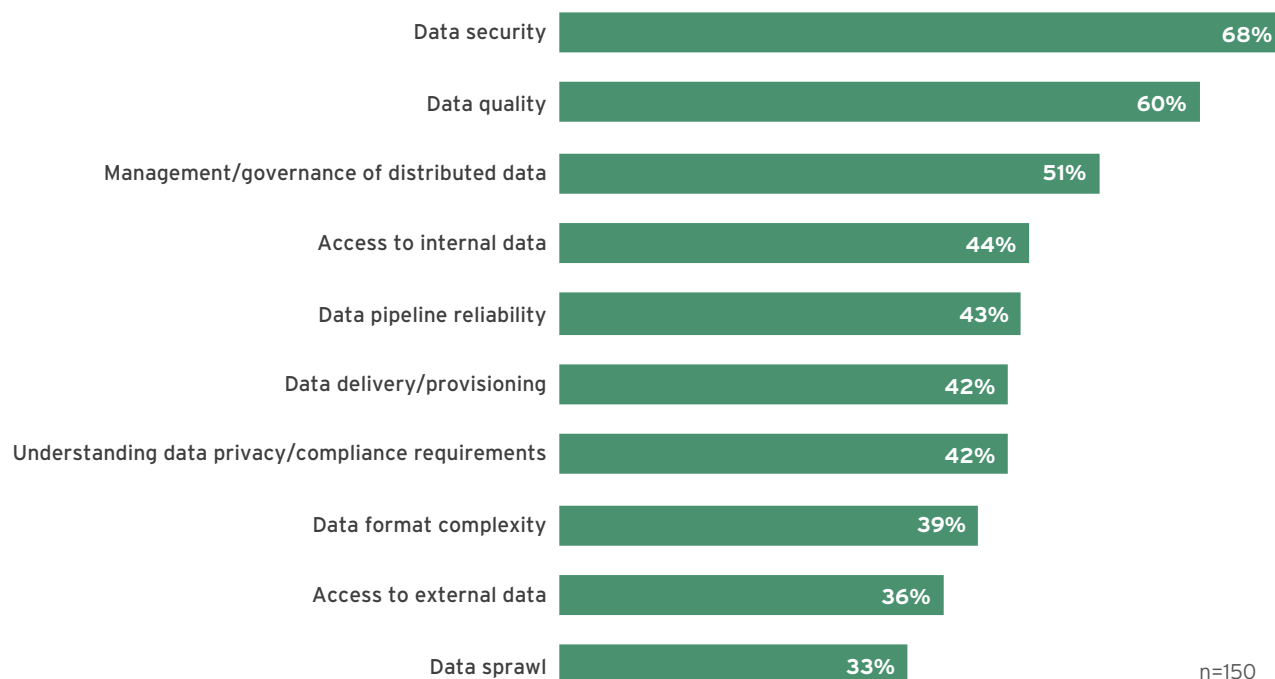
Data Management Challenges

One of the most significant challenges that organizations face as they seek to become more data-driven relates to the data itself: how to manage and govern it in an efficient manner to deliver the desired benefits. This is an age-old challenge, but one that is made more complicated by the increasing pace and volume at which data is generated.

Our DataOps Survey results indicate that data management is not a single challenge, per se, but encapsulates a range of challenges (see Figure 2). The most significant of these, according to our survey respondents, is security, which was cited by 68% of respondents as a data management challenge, followed by data quality (60%) and the management/governance of distributed data (51%).

Data security	68%
Data quality	60%
Management/governance of distributed data	51%
Access to internal data	44%
Data pipeline reliability	43%
Data delivery/provisioning	42%
Understanding data privacy/compliance requirements	42%
Data format complexity	39%
Access to external data	36%
Data sprawl	33%

n=150

Individually, each of these factors is a significant challenge. Data security is a perennial concern, of course, while addressing data quality is essential if organizations are to ensure they are making decision based on accurate data. Additionally, the fact that data is increasingly distributed across multiple data platforms (relational, nonrelational and Apache Hadoop/Spark), and multiple locations (on-premises and multiple clouds) complicates the understanding of what data organizations truly have at their disposal.

However, these factors do not occur in isolation. Survey respondents selected, on average, 4.7 different options as the primary data management challenges faced by their organizations. It is the confluence of this range of factors – including not just data security, quality and distributed management, but also data access, data provisioning, data privacy and data sprawl, along with reliability and data format complexity – that makes data management so challenging.

Never-Ending Data Growth

One of the primary trends shaping the data processing and analytics market in recent years has been the growth of data. This has driven the adoption of emerging data platforms and analytics techniques, and it has also been accelerated by the use of more economical data storage and processing products and services – which has enabled enterprises to store and process data that was previously ignored.

Our DataOps Survey results illustrate the continuing challenge in relation to the growth of data, the growing number of data sources, and a growing range of data processing locations (see Figure 3). In terms of data growth: 29% of respondents stated that their organization's data is growing at between 100-500GB per day, with 19% reporting growth between 500GB-1TB per day, 12% between 1TB-2TB per day, and 13% reporting data growth in excess of 2TB per day.

Figure 3: Data Growth Per Day
Source: 451 Research DataOps Survey

n=150
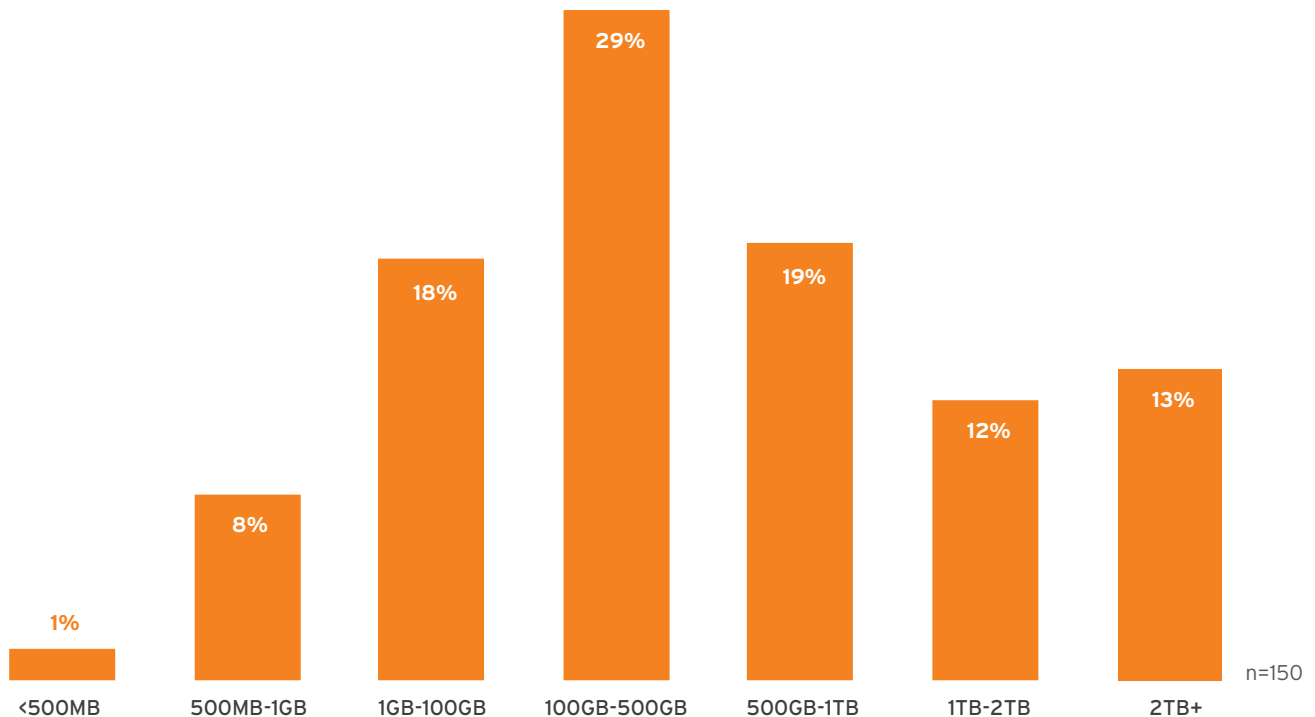
While traditionally enterprises had to deal with multiple relational database providers, perhaps with a few specialist nonrelational databases, the growth of NoSQL and Apache Hadoop means that many enterprises are now having to deal with a plethora of data platforms: 85% of respondents are using relational operational databases, and 72% have a data warehouse, while 60% are using NoSQL databases and 49% have data lakes based on Apache Hadoop/Spark.

In addition to the growing number of data platforms, the growing range of data processing locations is also a challenge, and one that is likely to grow as a greater volume of data processing and storage occurs in cloud computing platforms. The survey results indicate that a significant shift to the cloud is under way. While 41% of respondents say they are currently storing more than half of their data in the cloud, that is expected to increase to 73% of respondents two years from now.

The survey results indicate that the organizations that are most advanced in their use of DataOps are also more advanced in terms of data storage and processing in the cloud, which is perhaps not surprising given that the use of cloud databases is driven by a similar combination of factors, including the potential for operational efficiencies and greater agility.

The results indicate that half of respondents that consider themselves Optimized for DataOps are already storing and processing between 50% and 75% of their data in the cloud (compared with 30% of all respondents), while 53% expect to be storing and processing between 75% and 100% of data in the cloud in two years, compared with 38% of all respondents.

More Data, More Problems

While storing, processing and analyzing greater volumes of data brings potential advantages in terms of greater business intelligence, more data also means more data management challenges.

This is particularly true in relation to provisioning and managing data management environments: while agile development and DevOps practices mean that enterprises are able to develop and provision new application environments at a much faster rate than ever before, the provisioning of data to support those applications remains a bottleneck to operational efficiencies.

Two-thirds of survey respondents currently have four or more individuals involved in the data provisioning process, and those with the greater data growth require more participants: 42% of respondents with data growing at more than 1TB per day have more than five people involved in the data provisioning process, compared with 26% of all respondents.

Additionally, 78% of survey respondents indicated that it takes their organization four days or more to provision a new data environment, and those with the greatest data growth face greater delays: 21% of respondents with data growing more than 1TB per day take more than a week to provision a new data environment, compared to 14% of all respondents (see Figure 4).
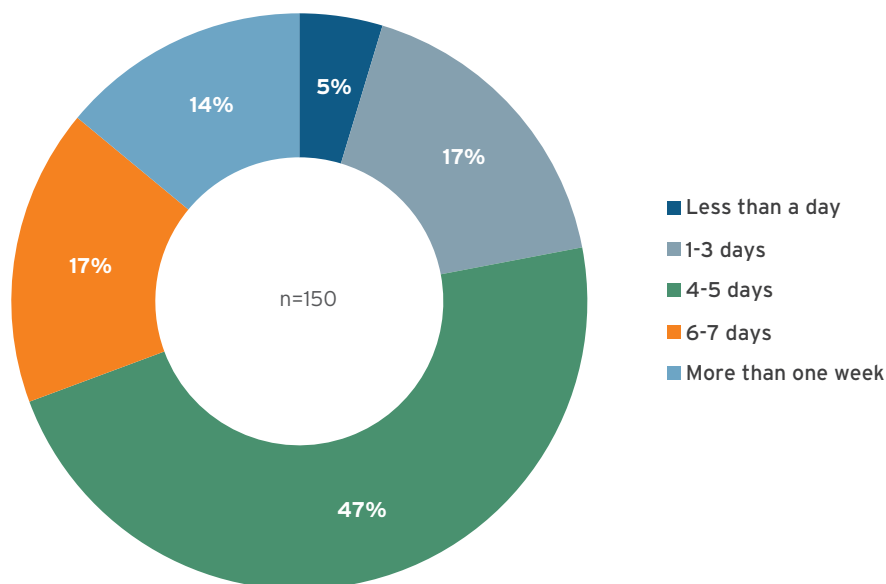
Complexity in the provisioning of data isn't just a data management headache: it also has significant implications in terms of business efficiency and productivity, as well as data security. According to the survey results, 65% of respondents agree (37%) or strongly agree (28%) that lack of access to the right data negatively impacts organizational efficiency, productivity and/or speed.

In particular this can lead to challenges with regard to sharing data – either internally with employees, or externally with customers and partners – and 68% of respondents agree (39%) or strongly agree (29%) that securing data shared with internal and external users is a concern within their organization.

How DataOps Can Help

There are indications that DataOps can help address these challenges. The DataOps Survey results show that the organizations that see themselves as Optimized for DataOps take less time to provision a new data environment, enabling greater productivity.

While DataOps is closely associated with agility and operational efficiency, the survey results also indicate that DataOps has the potential to help alleviate security and compliance as cross-functional concerns related to data management. In fact, concern about security and compliance is actually the number one driver for focusing on DataOps, cited by 65% of respondents.

While this is in keeping with security being the number one data management challenge (see 'Impact of DataOps' above), it is still interesting that security and compliance concerns score higher than improved operational efficiencies (63%), cost reduction/ROI (55%), the introduction of innovative, application-driven business capabilities (52%), and the need for greater business agility and speed (51%).

The fact that all of these responses were cited by more than half of survey respondents illustrates that DataOps is a multifaceted initiative that is by no means solely focused on agility. The list of priorities in terms of DataOps culture and processes tells a similar story. Again, security topped the list, specifically more secure and compliant processes (61%), while the smooth introduction of new tools and methodologies was cited by 60% of respondents, followed by greater collaboration and alignment across functions (57%), as well as organizational responsiveness and support for development (also 57%).

In relation to perceived business benefits of DataOps (see Figure 5), security is again top of mind with two-thirds of respondents citing greater security and compliance, followed by increased business agility and faster time to market (63%), as well as greater business insight and improved decision-making (55%), and the introduction of innovative, application-driven business capabilities (54%).

Figure 5: Business Benefits of DataOps

Source: 451 Research DataOps Survey

Benefit	Percentage
Greater security and compliance	66%
Increased business agility and faster time to market	63%
More business insight / improved decision making	55%
The introduction of innovative, application-driven business capabilities	54%
Cost reduction / operational efficiency	49%
Cross-functional alignment	42%

n=150
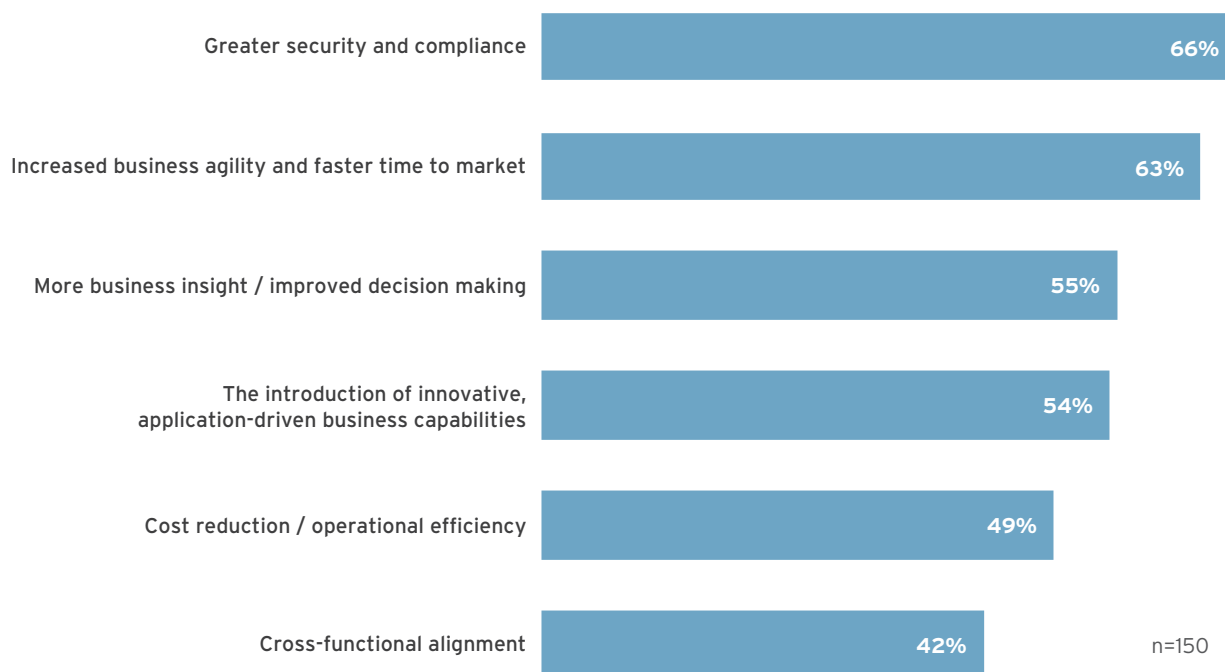
Another potential benefit of DataOps that wasn't explicitly addressed in the survey questions but is illustrated by the responses is reducing the potential for friction between the different groups of users involved in data management, including data operators and data consumers, as well as business executives.

The survey results indicate that the organizations with the greatest adoption of DataOps are more successful in involving people from a broader set of roles in using their database and data management products and services. Specifically, among representatives of organizations that consider themselves Optimized for DataOps, 75% see the application development group as primary users (compared with 46% of all respondents), while 81% see executives as primary users (compared with 54% of all respondents), and 81% see data/business analysts as primary users (compared with 63% of all respondents).

DataOps Traction

While DataOps is designed to reduce friction between data consumers (such as data analysts, developers and senior decision-makers) and data operators (including data management and IT professionals), it still starts with data operators. The fact that the companies that are most advanced in terms of their adoption of DataOps also see greater involvement in data management from people in multiple roles across the organization is evidence that internal adoption of DataOps practices is accelerating among those early adopters.

DataOps adoption is still in the early stages, although 451 Research expects it to have a growing impact on the market in the coming years. One of the reasons we are confident of that lies in the survey results indicating that DataOps is delivering on the drivers, priorities and benefits cited above.

As previously stated, almost all (92%) of survey respondents agree that improved DataOps would have a positive impact on their organization's success, with 45% strongly agreeing and 38% agreeing. The results also indicate that the organizations that are most heavily invested in DataOps have an elevated perception of its potential to deliver positive impacts. Of those organizations that are Optimized for DataOps, 97% agree that improved DataOps will have a positive impact on their organization's success, with two-thirds of these Optimized organizations strongly agreeing that improved DataOps will have a positive impact, compared with just 45% of all respondents.

Additionally, 89% of respondents to our survey expect to significantly or moderately change their processes, organizational structure or training/consulting to improve DataOps in the next 12 months, with 38% expecting significant change and 51% anticipating moderate change. Meanwhile, 86% of respondents plan to increase their spending, investment or development effort on DataOps technologies in the next 12 months, with 41% expecting a significant increase and 45% anticipating a slight increase.

Figure 6: DataOps Investment Plans
Source: 451 Research DataOps Survey

n=150

- Significant increase
- Slight increase
- Remain the same
- Slight decrease
- Significant decrease
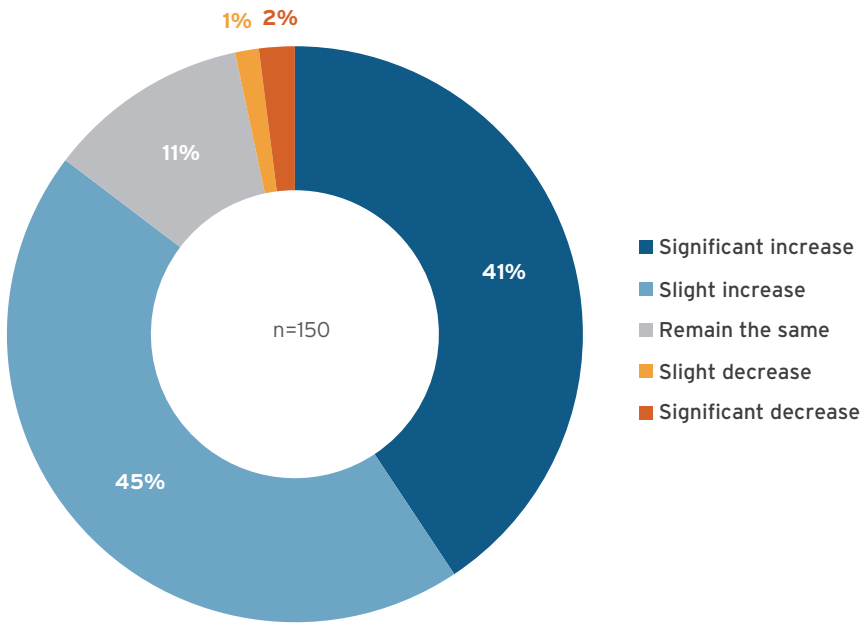
Again, the results also show that the greater the involvement in DataOps, the more likely an organization is to be increasing their engagement with DataOps in the future. Of those organizations that are Optimized for DataOps, 75% plan to significantly increase spending, investment, or development on DataOps technologies in the next 12 months, compared with 41% of all respondents.

Organizations that consider themselves Optimized for DataOps are also doubling down on DataOps-driven process/organizational change, with 69% of them expecting to significantly change their processes, organizational structure, or training/consulting to improve DataOps in the next 12 months, compared with 38% of all respondents.

The 451 Take

While the concept of DataOps is perhaps most associated with operational efficiencies, the survey results indicate that those efficiency improvements are not just related to agility, but also to security and transformational change. Companies that are already engaged with DataOps overwhelmingly agree that it is having a positive impact on their organization, and while improved agility and efficiency are closely associated with DataOps, the biggest driver, priority and benefit is actually related to security and compliance.

Those enterprises that have adopted DataOps are also more advanced in terms of transitioning to the cloud and executing digital transformation strategies – and as such they are better-placed to gain competitive advantage over their rivals. In addition, the early adopters of DataOps are enjoying benefits to the extent that they are doubling-down to invest even further in products and services as well as process and organizational change. Thus, the survey results reinforce our view that while it is still relatively unknown as a mainstream term today, DataOps can be expected to have a growing impact on the wider market in the coming years.
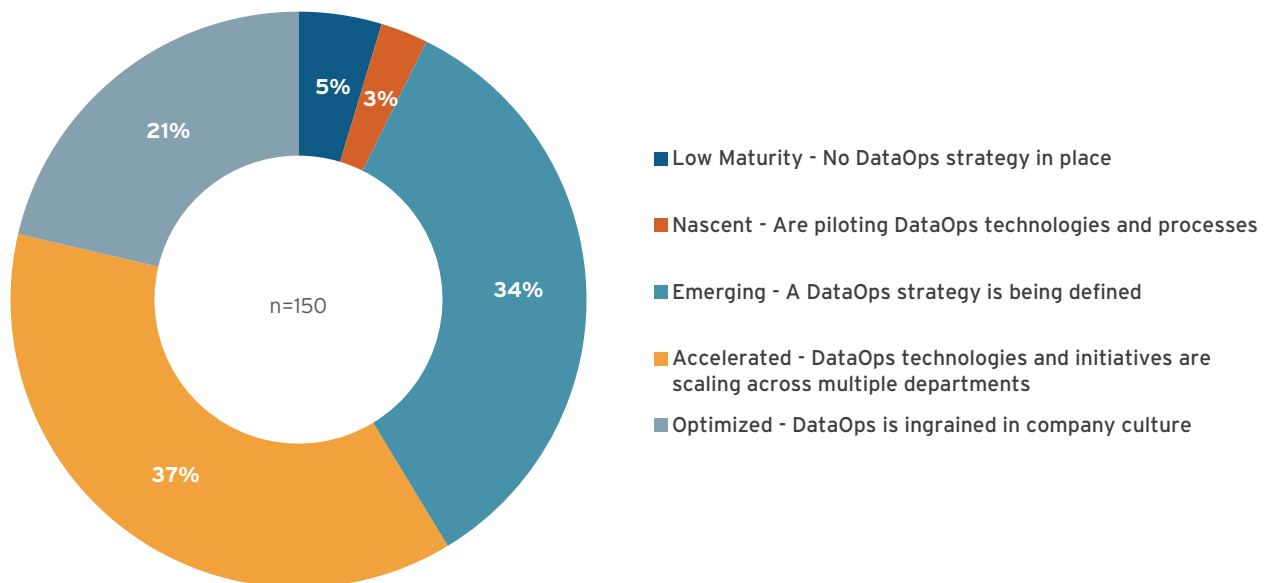
Survey Demographics

451 Research's DataOps Survey was conducted in October 2018 with 150 representatives of North American organizations with more than 1,000 employees and a minimum of 2PBs of data under management who had a solid understanding of their organization's data management strategy. Of these respondents, 39% were from companies with between 1,000 and 4,999 employees, 31% from companies with 5,000-9,999 employees, and 30% from companies with more than 10,000. Meanwhile, 27% of respondents represented companies with between 2PB and 5PB of data under management, 37% represented companies with 5PB-10PB, and 37% represented companies with more than 10PB.

A range of industries (banking and finance, healthcare, technology, manufacturing, retail, government/public sector, insurance and telecom/media) were included, as well as a range of roles (both junior and senior) covering development, testing, database management, IT operations and business intelligence. No more than 20% of respondents were involved in a single vertical industry, and no more than 20% represented a specific role.

Survey respondents were mostly mature in their level of DataOps adoption: 21% declared themselves Optimized (with DataOps ingrained in company culture), while 37% saw themselves as Accelerated (with DataOps technologies and initiatives across multiple departments), and 34% Emerging (in the process of defining a DataOps strategy). Only 3% considered their level of DataOps maturity to be Nascent (piloting DataOps technologies and processes), while 5% admitted to being of Low Maturity (with no DataOps strategy in place).

Figure 7: Level of DataOps Maturity
Source: 451 Research DataOps Survey

- Low Maturity - No DataOps strategy in place
- Nascent - Are piloting DataOps technologies and processes
- Emerging - A DataOps strategy is being defined
- Accelerated - DataOps technologies and initiatives are scaling across multiple departments
- Optimized - DataOps is ingrained in company culture

n=150

5% 3% 21% 34% 37%

About 451 Research

451 Research is a preeminent information technology research and advisory company. With a core focus on technology innovation and market disruption, we provide essential insight for leaders of the digital economy. More than 100 analysts and consultants deliver that insight via syndicated research, advisory services and live events to over 1,000 client organizations in North America, Europe and around the world. Founded in 2000 and headquartered in New York, 451 Research is a division of The 451 Group.

NEW YORK
1411 Broadway
New York, NY 10018
+1 212 505 3030

SAN FRANCISCO
140 Geary Street
San Francisco, CA 94108
+1 415 989 1555

LONDON
Paxton House
30, Artillery Lane
London, E1 7LS, UK
+44 (0) 203 929 5700

BOSTON
75-101 Federal Street
Boston, MA 02110
+1 617 598 7200