

The Data-Forward Enterprise:

How to Maximize Data Leverage for Better Business Outcomes

Phil Goodwin | Randy Perry

May 2020









IDC OPINION

IT leaders worldwide know that cloud is now an indispensable part of their IT infrastructure strategy. In fact, IDC research has found that 70% of CIOs have a stated cloud-first application deployment strategy. Most organizations will use a mix of private cloud and public cloud, referred to as hybrid cloud. In addition, we expect organizations to use multiple public cloud services, referred to as multicloud. Even if organizations are not currently operating in a multicloud environment, we expect that the vast majority will do so.

Valuable data is now commonly spread across multiple physical locations and different repository types. Data is stored in the core datacenters, in public cloud locations including SaaS applications, and at the edge, including remote locations, endpoint devices, and specialized IoT devices. In addition, this data may be stored as traditional structured data, widely varying unstructured data, newer NoSQL data types, and emerging container microservices. In fact, among the research findings for this study, determined organizations manage an average of 25 discrete data silos. These disparate silos, separated by either geography or technology, create a problem referred to as data sprawl or data fragmentation. This means that each silo must be managed for life-cycle policies, security and data protection policies, analytics, and so forth. Few organizations have an enterprisewide data management strategy and platform. Without the tools and platforms necessary to unify data across the enterprise, organizations cannot gain full competitive benefit from the data available to them. In addition, multiple silos complicate data security efforts and consume more human effort to manage, resulting in both monetary and lost opportunity costs.

Rubrik commissioned IDC to conduct an independent, worldwide primary research study to determine the magnitude of the data sprawl problem, how well IT organizations are prepared to deal with the problem, and how to create a Data Management Maturity Index for IT organizations to measure themselves compared with others in their industry.

Key findings from the study include:

-  **45.7%** of organizations believe they derive less than half of the potential value from their data due to **data management deficiencies**.
-  **Fewer than 33%** of organizations report having fully automated means for dealing with **data security, governance, backup, and disaster recovery (DR)**.
-  **More than 80%** of IT leaders identified **data sprawl** as a problem that must be addressed.
-  **More than 60%** of respondents rate managing multicloud, hybrid cloud, data security, and disaster recovery as **“major” or “extreme” challenges**.
-  **Only 9.2%** of organizations have **a single, centralized data management** system or platform.
-  **Investing in cloud computing** — private, public, hybrid, and multicloud — is **the highest priority** for organizations, indicating that they will be managing data across the spectrum of cloud deployments.

For those organizations with more mature enterprise data management practices, the results were tangible:

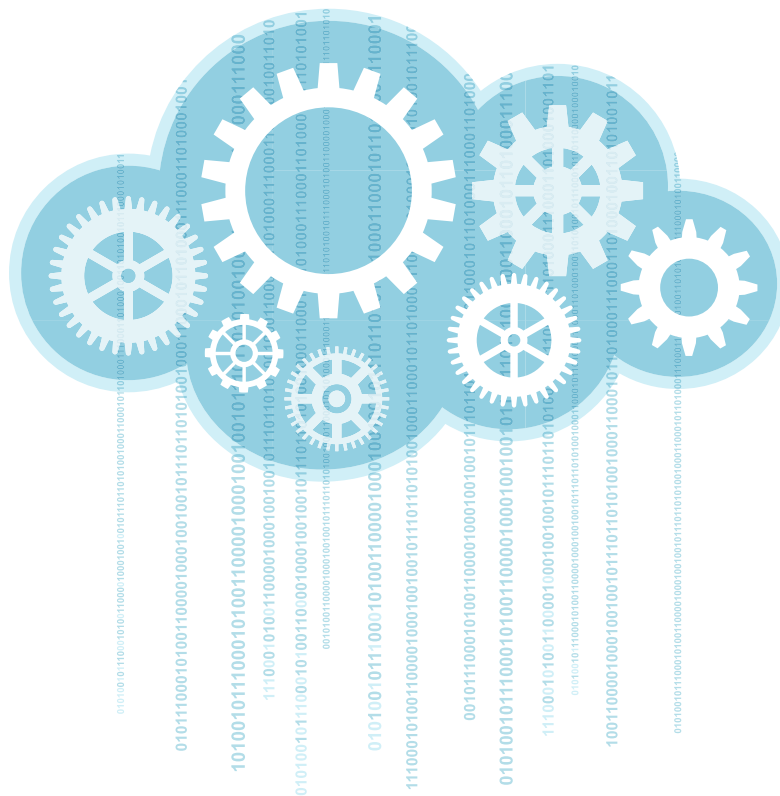


IDC believes that an enterprisewide data management system or platform is essential for “data driven” organizations to gain or maintain competitive advantage in today’s digital economy.

METHODOLOGY

This study involved a primary research survey of 801 respondents with statistically relevant sample sizes in North America, Western Europe, and Asia/Pacific. The survey respondents represented a wide range of industries from small and medium-sized businesses to large global enterprises. All respondents had either direct responsibility for, or direct knowledge of, the organization's data management practices; 60% were senior-level IT professionals and 40% were senior line-of-business (LOB) leaders.

In addition to the primary research instrument, IDC conducted in-depth cross-industry interviews with three IT leaders and two LOB leaders, three from North America and two from Europe. These wide-ranging interviews allowed for detailed exploration of an organization's data management practices and quantifiable results. Interviewees were unaffiliated and unknown to the study's sponsor and were entirely anonymous.





RESEARCH FINDINGS

Multicloud data control is emerging as a serious and mainstream initiative for IT and business leaders alike. IDC research indicates that 60% of organizations have either completed or started digital transformation (DX) initiatives. The objective of these initiatives is to move the organization toward being data driven, whereby the advanced use of data management and analytics helps the company derive a competitive advantage. Yet data sprawl — that is, data stored in different locations and in different formats and derived from different applications — makes the aggregation and correlation of data across the enterprise challenging. If organizations cannot aggregate this data so that it can be analyzed meaningfully, then the organization cannot utilize it fully.

Data stored and managed in this manner results in data silos that increase data management complexity. Silos initiate unnecessary human effort and redundant management products useful for only that siloed environment. They also introduce additional risk of data loss with additional malware vulnerabilities and more challenging DR. The bottom line is reduced data visibility and loss of data control. IT leaders are now recognizing the need for, and challenges of, multicloud data control.

These challenges were expressed during an in-depth interview with the CIO of a large United States-based technology company. This individual spoke to the organization's challenge in dealing with data sprawl. This organization is compelled to spend a significant amount of time simply identifying and classifying data. Lack of controlled visibility into its fragmented data increased the risk of potential regulatory and compliance violations. This individual attributed much of the problem to the lack of a data management platform, describing the current system as just a collection of tools, impeding business and operational agility. One key desire for this company is to have a unified data management platform across the enterprise in order to reduce the time and manual effort needed to manage data.

Several factors are emerging, which combined make multicloud data control a particular challenge:



The first factor is accelerating data growth. Previous estimates for data growth have been in the range of 30–35% compound annual growth rate (CAGR), whereas the results from this survey showed that the respondents expect a 46% CAGR on average. At this anticipated rate, data will more than double every two years. The sheer volume of data, while not unprecedented, will strain both infrastructure and data management processes.

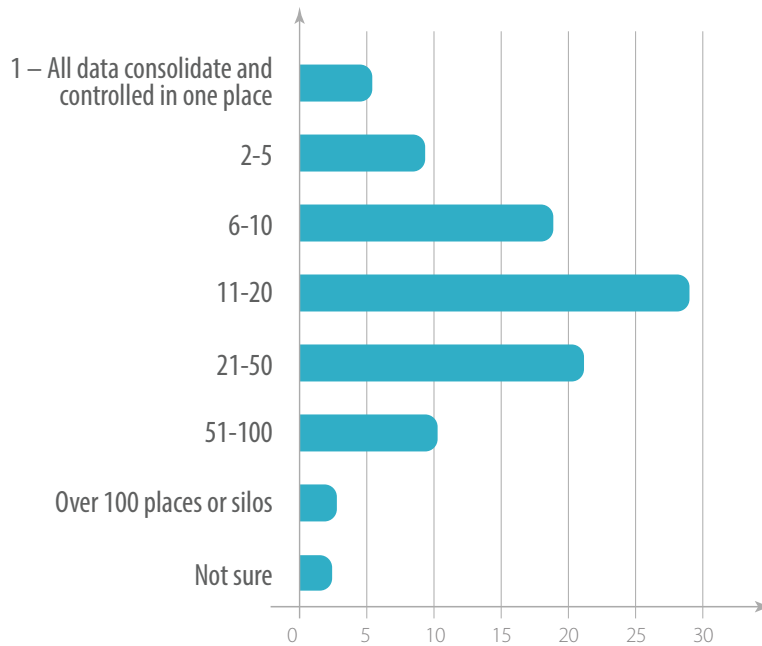


Second, surveyed organizations expect their data to be almost evenly distributed across multiple public clouds, with 26% of data, and on-premises private cloud, off-premises private cloud, and traditional on premises, each with 22% of data (the remainder is other/don't know). When we compare these results with those of prior surveys, we believe this shows a trend away from traditional on premises to clouds of all types.



Third, data silos are significant and proliferating. Data from this survey showed that organizations average 23 separate data silos per organization (see Figure 1). Each silo must be separately managed with its own data policies, including security, protection, governance, and life-cycle management. This management effort often involves different tools and processes that require redundant purchases, training, and maintenance.

Figure 1. Number of Data Silos per Organization



n = 801 Base = all respondents Average = 23 Source: IDC's Rubrik Thought Leadership Study, December 2019

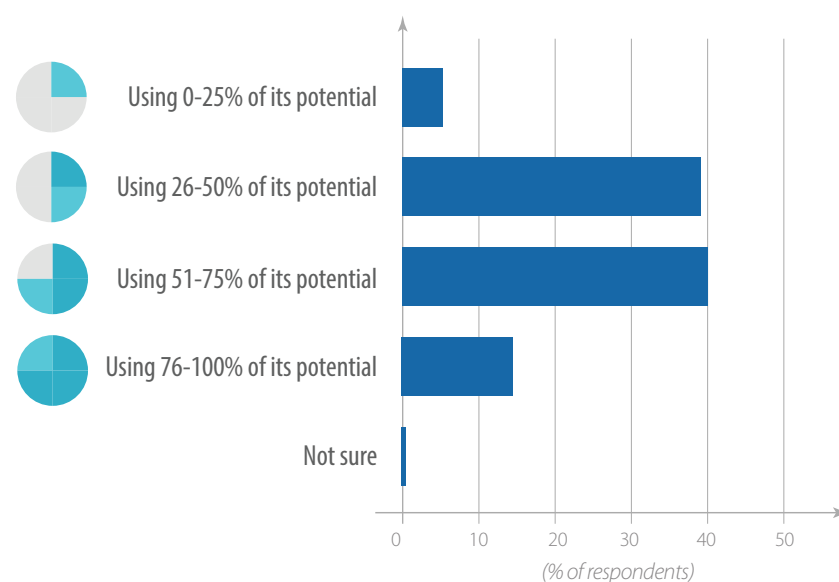
Notes: The survey is managed by IDC's Quantitative Research Group. Data is not weighted. Use caution when interpreting small sample sizes.

Forward-looking organizations are constantly seeking ways to improve their use and leverage of data. This point was emphasized by another of our in-depth interviews, this time with a senior IT executive of a large United States–based financial institution. A key objective for this organization is to provide better service to customers. Leveraging data and using advanced data analytics will help the organization identify which customers to focus on and the specific needs of those customers. The top priority is better information so that the organization can make faster, more accurate decisions. However, this executive also noted that the organization is challenged by data scattered across the organization that cannot be easily discovered or classified because there is not a unified data management solution across the enterprise. Moreover, the organization is forced to make tactical decisions and deploy new functionality at the expense of optimizing existing systems. The organization plans to use cloud computing to improve its agility and scalability but realizes that this is going to escalate the challenges around data sprawl.

The CIO of the technology company previously mentioned offered an illustration of the value that can be achieved by leveraging existing data. The company was able to reduce customer churn from 20% to just 7% annually by using analytics to predict customer behavior and then taking proactive engagement action.

Across industries, our survey found that organizations frequently are unable to fully leverage the potential of data available to them. In fact, our survey found that 45.7% of organizations believe they leverage less than half of the value of their corporate data (see Figure 2).

Figure 2. Degree of Data Leverage



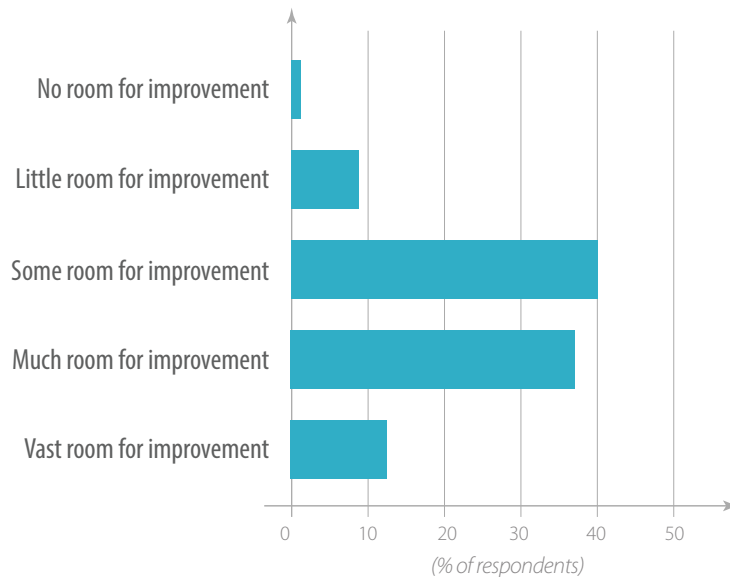
n = 801 Base = all respondents Source: IDC's Rubrik Thought Leadership Study, December 2019

Notes: The survey is managed by IDC's Quantitative Research Group. Data is not weighted. Use caution when interpreting small sample sizes.

From Figure 2, we can also see that only 14.7% of organizations reported achieving a 76% or higher value from available data. We believe that the inability to fully leverage the value of data is a direct result of the challenges in correlating data across data silos, as corroborated during another in-depth interview with a senior IT director of a large United States-based manufacturing company. According to this individual, the IT organization's number 1 challenge is dealing with data sprawl across applications, platforms, and array formats. The organization's key aim is to synchronize data across on-premises and cloud repositories.

Although cloud is certainly a means of improving application agility, data management, and control, fully utilizing cloud computing entails its own set of challenges. From our survey, a whopping 90% of respondents — representing a preponderance of both IT and LOB leaders — believe they have room to improve their data management practices. In fact, 49.9% of respondents said they have “much” or “vast” room for improvement (see Figure 3).

Figure 3. Room for Improvement in Data Management



n = 801 Base = all respondents Source: IDC's Rubrik Thought Leadership Study, December 2019

Notes: The survey is managed by IDC's Quantitative Research Group. Data is not weighted.

Use caution when interpreting small sample sizes.

We asked survey respondents to tell us what their top priorities for data management were. Data security was mentioned as the number 1 priority and cited by 37.4% of respondents. Other top priorities cited included management of data-intensive applications (33.3%) and dealing with increasing regulatory requirements (33.3%). The greatest single data management challenge cited by respondents was dealing with the increasing number of locations where data is stored (18.2%). We believe that the combination of data growth and an increase in data locations, as well as compliance requirements plus the frequency and severity of malware and ransomware attacks add up to major data security challenges for organizations. The security issue was highlighted during our in-depth interview with the CIO of the large technology company who stated that data sprawl causes problems with data visibility and security for HIPAA, GDPR, and the pending CCPA (California Consumer Protection Act). This individual summarized the problem by stating, "... *data management is a huge exercise for us. It's a holistic picture — where do I have it, how do I manage it, what is it's life cycle, who has access to it, and how do I secure it.*"



DATA MANAGEMENT MATURITY INDEX

In recognition of the size and complexity entailed in data management, Rubrik and IDC partnered to conduct research about current and future data management practices, adoption, and maturity in organizations. The analysis of this research effort is the creation of a Data Management Maturity Index. The Data Management Maturity Index:



- Shows where you are on your data management journey
- Enables you to benchmark how you compare to your peers







- Provides guidance on what steps to take to advance to the next level
- Measures business improvements resulting from higher levels of data management maturity



Description of the Methodology for Developing the Index

The methodology for developing the Data Management Innovation Index is designed to segment the full body of research into distinct subgroups based on how organizations approach data management and how their approaches correspond to improving business outcomes.

The data analysis identified four distinct data management levels of innovation:

	Leading top 10%	Leading innovators are aggressively using data to disrupt new markets. With over 76% annual growth, they are effectively leveraging more than 75% of their data and have already implemented the practices, policies, and technology to include a single unified data management system.
	Advanced 60-90%	In advanced innovators, business and IT goals are aligned at the enterprise level around the need to leverage and manage their quickly growing data. They have made investments in technology, have implemented practices and policies to improve data management, and have begun to deal with data sprawl. They have also evaluated a single unified data management system.
	Emerging 30-60%	For emerging innovators, LOBs are becoming more dependent on data and realize the need to develop a data-driven business strategy. Leveraging data is still not effective, but these organizations are considering investments to deal with data sprawl.
	Experimental 30% bottom	Experimental innovators' business objectives are not yet dependent on data initiatives. They are not effectively leveraging their data; they are not even aware of data sprawl, and their data management is poor.

Key Attributes of Leading Innovators

For leading innovators, the importance of data is reflected in organizational culture. Leading innovators are exploiting data to drive revenue, reduce operational costs, and improve profit.

These organizations are using data to improve customer satisfaction and retention by developing and delivering new products faster. They are using data insights to improve top-line revenue by launching new products and services offerings and acquiring new customers. They are better at mitigating risk and more strongly adhering to security and compliance regulations. In detail:



Size of current data environment. Leading innovators have two to six times the data of their peer groups.



Annual growth of data environment. Leading innovators' data is growing 50–100% faster than other groups.



Extent to which data is leveraged. Leading innovators are heavily leveraging their data to its potential at a rate of 4–10 times the rate that experimental innovators do.



Sufficiency of technology in place for data management and control. 89% of leading innovators feel they have all the technology in place to address key organizational challenges to data management, while only 7% of experimental innovators have all the technology in place.



Approach to data sprawl. 82% of leading innovators have identified data sprawl as a top priority and 51% have implemented a strategy to address it. Less than 1% of experimental innovators have identified data sprawl as a top priority and 2% have implemented a strategy to address it.



Policies and practices in place to address data management and control. Over 50% of leading innovators have implemented the 6-9 initiatives necessary to control data (see the Business Value Results section). Less than 30% of experimental innovators are in agreement as to which actions are necessary to control data.



Investments and plans for a single unified data management system. At 64%, leading innovators are two times as frequently organized around one centralized group for data management and 86% have already invested in a single unified and consolidated technology system for the control and management of data. Less than 2% of experimental innovators have invested in such a system.

Table 1 outlines characteristics of organizations at various stages of the Data Management Maturity Index.

Table 1. Data Management Innovation Index

	Experimental	Emerging	Advanced	Leading
Amount of data (>50PB) (%)	22	31	42	66
Annual data growth (%)	37	41	52	76
Annual application growth (%)	26	31	39	62
Levering potential from all data (%)	<50	50	>50	>75
Management technology in place	50% most or all	79% most or all	92% most or all	100% most or all
Familiarity with data sprawl	Least	Less	More	Most
Addressing data sprawl	Not considering	Considering	Started on	Implemented
IT policies/practices in place to improve data management	Least	Less	More	Most
Single unified data management system	Researching	Considering	Evaluating	Invested

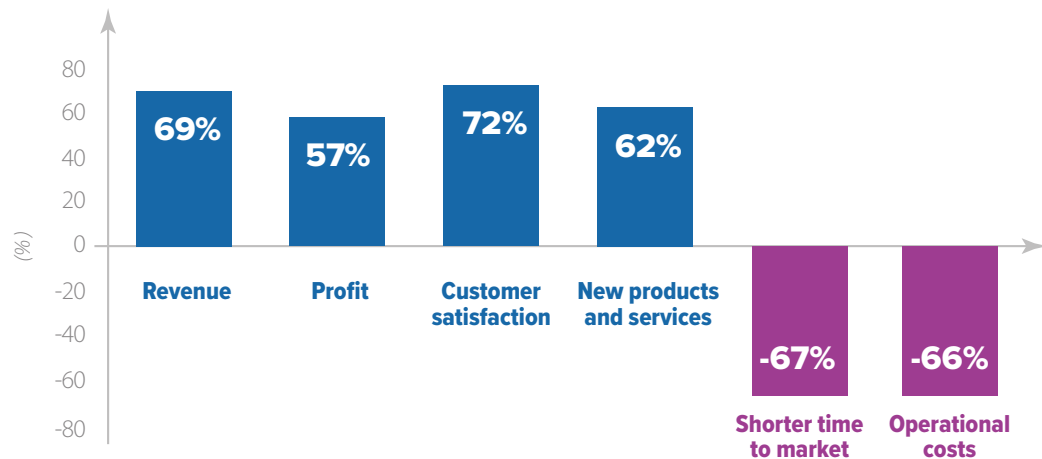
n = 801 Base = all respondents Source: IDC's Rubrik Thought Leadership Study, December 2019

Notes: The survey is managed by IDC's Quantitative Research Group. Data is not weighted. Use caution when interpreting small sample sizes.

Business Outcomes

The most innovative organizations at the leading innovator level experience two to three times improvements to business outcomes when compared with those organizations at the experimental innovator level. These organizations also enjoyed significant business performance gains (see Figure 4).

Figure 4. Business Outcome Improvements of Leading Innovators



n = 801 Base = all respondents Source: IDC's Rubrik Thought Leadership Study, December 2019

Notes: The survey is managed by IDC's Quantitative Research Group. Data is not weighted. Use caution when interpreting small sample sizes.

Improving Data Management Maturity — Next Steps

Readers can do assessments of their own organization's data management maturity and benchmark where they are relative to the innovation levels identified in the index. In doing so, readers can identify areas for improvement, regardless of where their organization is benchmarked.



Experimental

Obviously, there is room for improvement, but your organization may not be ready to make the investments needed nor know exactly how to get going. Here are some initial steps to get started:

- Begin with a frank assessment and communicate that assessment with senior business leaders:

- Consider a peer analysis, as data management maturity should be assessed within the context of industry peers.
- Identify and quantify the impact of missed business opportunities.
- Conduct gap analysis in three distinct areas:
 - **Technology.** Identify what key technologies the organization is missing, such as a data management platform, data ingest, and data analytics.
 - **Process.** Examine what processes are needed to gather and manage data.
 - **People.** Identify the people who should be consuming the data and who should be providing it, as well as whether or not the organization has the necessary skill set to effectively create a data management capability.



Emerging

Most companies are in this stage when they first recognize the need for change and to move toward new strategies. If you are in the “emerging” stage and you are not already evaluating next steps and looking to make data management improvements, you are already behind. Some things to consider are:

- Evaluate enterprisewide capabilities. Many emerging organizations have some data control capabilities, but these efforts may be fragmented, siloed, and incomplete.
- A data management SWOT (strengths, weakness, opportunities, threats) analysis can help identify areas of strength to build upon as well as those areas that should be prioritized (i.e., threats containing competitive risk, then weaknesses, and then opportunities).
- Gap analysis includes assessing people, process, and technology aspects.
- Success criteria includes defining what multicloud data control success looks like for your organization (i.e., profit improvement and time-to-market improvement).



Advanced

Organizations understand the need for better data management, and they may have begun to implement the data initiatives and define the criteria for success. They are likely aware of and have addressed many of their technology needs and are actively consolidating multicloud data

management platforms. However, for these organizations, the mostly likely gap is in people and process:

- Consider establishing data architect and data manager roles for DataOps capabilities with the charter to create a bridge between IT capabilities and business needs to most effectively connect data providers with data consumers.

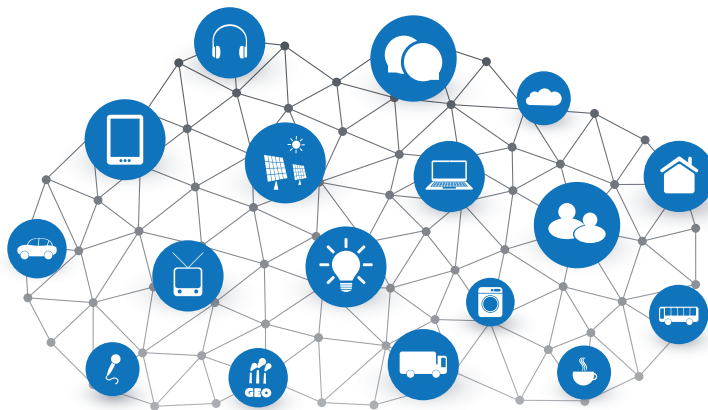


Leading

Even organizations that are in the top 10% still need to be forward thinking and to acknowledge the fact that data will only become more complex and challenging to manage over the next few years.

These organizations should:

- Establish a continual feedback loop among senior business leaders, DataOps teams, and evaluation criteria.
- Update people (i.e., training), process, and technology.
- Look for unique ongoing opportunities for competitive advantage.





BUSINESS VALUE RESULTS

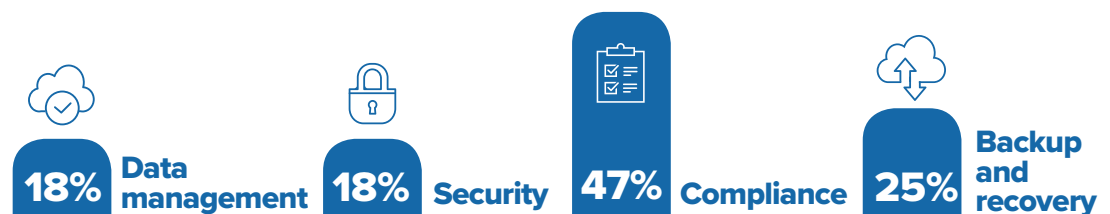
In addition to the quantitative survey, IDC conducted five in-depth interviews with organizations considered to be leading innovators. Three of these interviews were with IT leaders in which we focused on quantifying the impact of data management initiatives on their IT operations and uncovering how they are leveraging data to turn challenges into opportunities. The following initiatives were explored:

- Consolidating all data onto a single data platform
- Deploying data policy engines
- Deploying metadata management solutions
- Automating workflows for monitoring data
- Automating workflows for service delivery
- Implementing an agile DevOps approach to software development
- Standardizing policies for data governance across the organization
- Creating a centralized group dedicated to the control and management of data
- Addressing data sovereignty issues



The organizations implementing data management initiatives saw compelling benefits.

- **Increased IT operations productivity.** By enabling the IT staff to better manage data operations to scale, overall staff productivity increased by 24%, or the equivalent of 11.8 FTEs. As a result, IT staff is spending 64% more time on new initiatives rather than “just keeping the lights on.” Elements of increased IT operations productivity include:



- **Reduced business risk.** Data management initiatives provided a more reliable operating environment. As a result, unplanned downtime decreased 59%, which added 6.9 hours of productive time to every employee and returned an average of 0.5% of revenue to the organization. In addition, organizations were able to avoid or mitigate any regulatory or compliance breaches, lowering related fines by 90%.

- **Increased business growth.** Better management and control of data increased DevOps productivity, yielding a 68% increase in application development. Organizations increased the number of new business applications developed by 116% and increased the number of new applications deployed to the cloud by 289%.

Figure 5 illustrates the breakdown of annual average benefits.

Figure 5. Average Annual Benefits from Adopting Data Management Initiatives



*Note: Data is from five in-depth interviews with leading innovators.
Source: IDC's Rubrik Thought Leadership Study, December 2019*



FUTURE OUTLOOK

We believe that data sprawl will impact nearly every organization and complicate the ability for organizations to effectively use data. IDC forecasts that as many applications will be deployed in the next five years as have been deployed in the previous 40 years. Our research also shows that 80% of new applications will be deployed outside of the traditional datacenter. This shift will impact small and large organizations alike, as all organizations will increasingly rely on SaaS applications where the data itself is managed by a third-party provider. As such, secondary access to the data for analytics and other purposes will be more challenging and less subject to IT data management standards.



Five key elements are essential to establishing a multicloud data management and control capability:

- **Consolidated platform.** IT organizations will need a single platform in which to gather, consolidate, and federate information from disparate sources. In many cases, this will result in a data lake or similar deployment where analytics can be run.
- **Centralized policy engine.** Proactively, IT managers will want to manage the life cycle of data according to consistent policies based on data type and governance requirements.
- **Metadata management.** Since data has many sources and formats, metadata management will be an essential capability for data control. Using a metadata approach to data management

also enables organizations to extend the value of the data to drive assurance, security, governance, and disaster recovery, regardless of location.

- **API-driven extensibility.** The cloud has become integral to modernizing the IT environment. And when it comes to cloud, APIs deliver extensibility, automation, and self-service, vastly increasing agility and operational consistency. IT organizations must look for extensible frameworks that replicate what is available in the cloud, including trends like continuous integration and continuous delivery (CI/CD). Enabling DevOps, with the ability to build, integrate, and test via automation, reduces repetitive manual inefficiencies and opportunities for mistakes.
- **Search, ingest, and classify.** The consolidated platform must have the capability to search, ingest, and classify data. This capability will be largely enabled by metadata management capabilities and will be essential to life-cycle management to determine where data is, what the data is, who owns the data, and what policies should apply to the data from creation to destruction.



From the perspective of those on the front line, the following quotes from our in-depth interviews capture the consensus sentiment:

- *"The data management strategy is the key to digital transformation for us. Enabling the business to operate at a digital level, getting away from legacy systems, and having current data available at all times."* — **Senior IT Director, Manufacturing, United States**
- *"We really need a tool or technology that would allow us to get visibility to the data that exists in our environment, visibility to the age of that data and, ideally, some capability to associate the data with the different [data] stores."* — **SVP, Financial Services, United States**
- *"[Data management] allows us to focus on research and development, which is key for our business in order to continue innovation over the course of time."* — **Head of Strategy, Manufacturing, European Union**
- *"What would be great is if there would be a more open and common ecosystem [compared to proprietary AWS or Azure]."* — **CIO, Technology, United States**
- *"It's all about risk. It's about understanding if there is a data infringement or data security issue, [and] what does it do to our organization. Measuring the data risk and that's how we quantify it [data management]."* — **Senior Business Manager, Financial Services, United Kingdom**



CONCLUSION

Data sprawl is an emerging and serious problem for IT organizations. IDC research shows that digital enterprises (i.e., organizations with formal processes for using data as a competitive advantage) have eight times the growth rate and double the profitability of peers in their industries that are not digital. We believe many leading innovators could also qualify as “digital enterprises.” Obviously, those organizations that cannot effectively leverage their available data cannot achieve superior business results. The survey conducted for this paper shows that IT leaders recognize the evolving challenges around data sprawl, management, and control, including the inability to fully utilize the value of data, manage the number of data silos, and achieve the full value of cloud computing.

Our analysis reveals the value of modernizing data management and control that spans multiple clouds and hybrid architectures. Organizations with a strong data management strategy and control capabilities achieve tangible business results in terms of better employee productivity, faster time to market for new applications, IT agility and business resilience, and better use of the data spanning multiple clouds. Organizations can start on the path to better business outcomes by candidly assessing their current data management capabilities highlighted in this paper. IT leaders should also investigate multicloud data control platforms with the appropriate data policy engines and data ingest and life-cycle management capabilities. Organizations seeking to gain competitive advantage will take appropriate action immediately as those organizations that move faster and more effectively to solve for data control are much more likely to thrive in the competitive marketplace.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.



IDC Research, Inc.

5 Speen Street

Framingham, MA 01701

USA

508.872.8200

Twitter: @IDC

idc-community.com

www.idc.com

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2020 IDC. Reproduction without written permission is completely forbidden

