Value, Risk and Potential

A Modern Approach to Managing Data

Pathfinder

November 2021

Commissioned by

DATA DYNAMICS®
Manage Your Data®

451 Research

S&P Global
Market Intelligence

About this paper

A Pathfinder paper navigates decision-makers through the issues surrounding a specific technology or business case, explores the business value of adoption, and recommends the range of considerations and concrete next steps in the decision-making process.

About the Author

Paige Bartley

Senior Research Analyst, Data Management

Paige is a Senior Research Analyst for the Data, AI & Analytics Channel at 451 Research, a part of S&P Global Market Intelligence, covering core data management topics. Key themes include data privacy, data governance, data integration, metadata management, data quality and master data management. She has experience covering a broad range of technologies spanning the Data, AI & Analytics stack, from databases to self-service analytics.

In her current research, Paige is analyzing the need for information governance to maximize the value of enterprise data amid proliferating global regulatory requirements and rising consumer expectations for data stewardship. With data privacy and compliance as a specialty focus, Paige explores how the enterprise can align technical requirements with business strategy, enabling more profitable and compliant leverage of data.

Early in her career, Paige worked on the vendor side, providing marketing and strategy for ZL Technologies, an information governance provider that specializes in the management of unstructured data for compliance, legal and archiving needs. Prior to working at 451 Research, she was a Senior Analyst at Ovum (now Omdia).

Paige received her Bachelor of Science degree in psychology and neuroscience from Duke University, and a Master of Management Studies (MMS) degree from the Duke University Fuqua School of Business.

Executive Summary

There should be little surprise that data is increasing in importance to organizations as they attempt to become more data-driven in strategic decision-making. As global business conditions over the past couple of years have demonstrated well, critical shifts in consumer behavior have created more digital data and content than ever before, and potentially difficult competitive and economic conditions have underscored the importance of making the correct decisions swiftly to remain viable as a business. However, higher volumes of data and increased pressure to derive useful insight create internal pressures and challenges of their own. Many barriers remain for businesses to become more data-driven, and solving these challenges – whether they be cultural or technological – is a key to success.

In an era where data is becoming an enterprise-wide asset, organizations need to consider the full spectrum of data stakeholders and the stack of supporting technologies that they depend on, from data storage all the way up to cloud strategy and end-user applications. Every interaction with data presents a potential to derive business insight, but only if the data itself is managed effectively and consistently throughout the IT environment. Issues like duplicate copies of data can muddy the water for insight and productivity initiatives, and while it is likely unreasonable today to seek a single monolithic view of data assets, systematically managing the effects of these issues can be invaluable.

Aligning business strategy and objectives within the organization is also important today. Many business functions have different motives, uses for data and technology tooling preferences. Differences in cloud preferences and adoption highlight these internal divides and can slow high-level efforts to formulate an effective data strategy. But as the enterprise strives to build a functioning 'data culture' where all relevant workers can leverage data to drive business value, overarching agreement becomes key. Focus on cloud compatibility, data governance strategy and proper supporting technology can help these efforts.

Key Findings

- Data is growing in importance for organizations today. About 90% of enterprise survey respondents agree that data will be more important to their organization's decision-making one year from now.
- Barriers to being more data-driven can be either cultural or technological in nature. Top technological challenges include data privacy and security, as well as integration with legacy systems.
- Organizations actively seek to generate insight from data more quickly, with more than 45% of survey respondents identifying 'business agility improvements' as a key desirable benefit.
- While storage costs have declined over time, data volume has accelerated. Data/capacity growth remains the top storage pain point, with 36% of survey respondents in agreement.
- Multiple copies of business data present a data management challenge. An average organization has 5.4 copies of an individual piece of business data.
- Ideological differences can exist between business functions when it comes to storage. Nearly half (49%) of DevOps survey respondents reported 'shifting to the cloud' as a priority, while nearly one in five (19%) of storage-affiliated respondents reported their organization archives *none* of its data on public cloud architecture.
- Cloud compatibility, today, is a critical consideration; 72% agree that the ability to run data management/analytics software on multiple cloud/datacenter environments is an important consideration for their organization when selecting a new data management/analytics vendor.
- Data governance efforts are delivering value in proactive ways. Today, 38% of enterprise practitioners indicate 'faster access to relevant data' as a benefit delivered by governance.

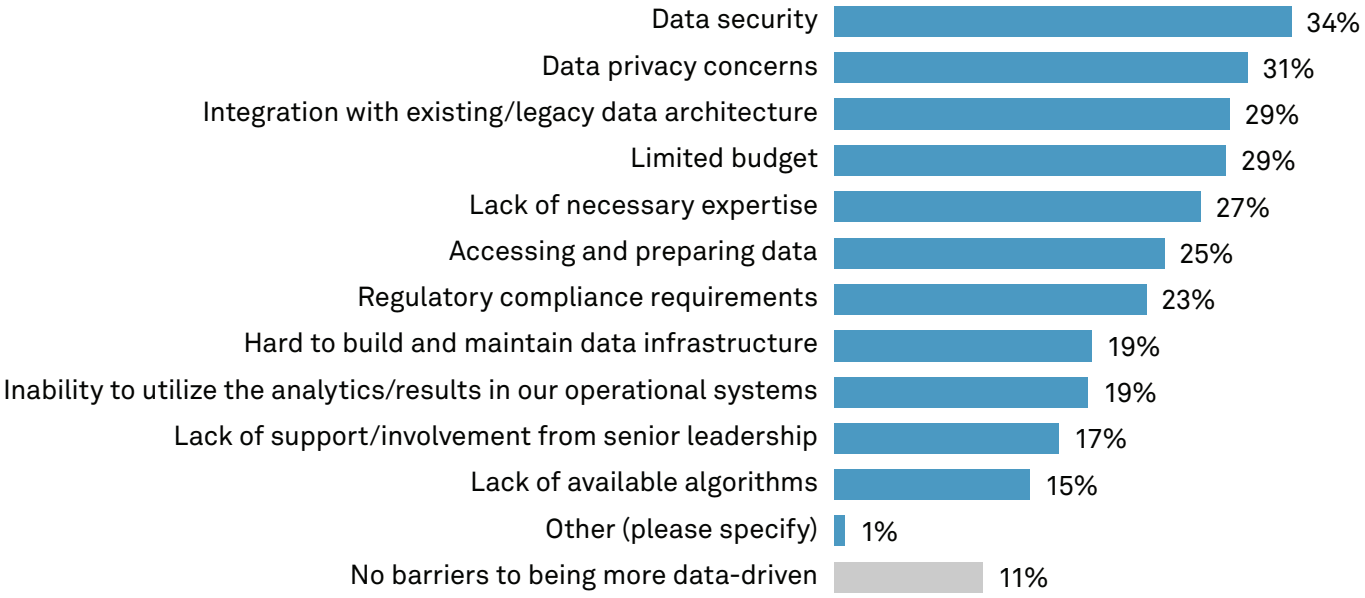
Technological Discussion

The Drive to Become a Data-Driven Organization

For organizations, there is typically little doubt today that data is gradually increasing in importance for business strategy and that the technology supporting the management of data and content is critical for ongoing progress in becoming more data-driven. Indeed, trends suggest that data is becoming substantially more important than ever. Based on 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2021 survey, 90% of respondents agree that data will be more important to their organization's decision-making one year from now. That response is up from 81% agreement when the same question was asked in a survey about six months earlier. But data itself is of limited utility when the structure and supporting elements that surround it – infrastructure, platforms, software, etc. – are not well coordinated. People and processes are also critical to success with data. The challenge spans the dynamics of the enterprise.

Amid this rise in data importance, businesses are looking to clearly establish an organizational data culture in which all relevant workers are empowered to leverage relevant information in their daily roles to drive better business decisions. But there do remain challenges in the attempt to become more data-driven. Limiting factors to success, perhaps unsurprisingly, are not simply technological limitations; challenges can be cultural as well.

Figure 1: Barriers Faced in Attempting To Be more Data-Driven



Q. What are the most significant barriers your organization faces in attempting to be more data-driven? Please select all that apply.
Base: All respondents (n=371)
Source: 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2021

Across organizations, many of the top barriers today are, indeed, technological in nature, but that should not depreciate the cultural concerns such as budget and lack of expertise that remain apparent today. There are many challenges here, and many of the challenges are closely clustered. However, when technological issues are isolated, key themes emerge related to today's distributed and heterogeneous – simply put, complicated – IT ecosystems.

Integration with legacy architecture continues to be a major challenge because it is typically not feasible or economically viable to just rip and replace an existing investment from the IT environment. Data privacy and data security issues – consistently at the top of concerns today – are salient pain points, visible especially acutely to the most data-driven organizations. The more an organization strives to leverage data, the more evident any perceived restrictions become. Highly data-driven organizations also tend to have more diversified IT architecture, complicating the consistent controls needed for compliance.

Overcoming these challenges requires more consistent management of all data, not just traditional structured forms of data that have historically existed in relational database systems. If organizations can attain consistent management and control, the benefits to be reaped are numerous, ranging from business efficiency improvements, to driving entirely new value, to the ability to respond more swiftly to external forces.

When 451 Research participants in the Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2021 survey were asked what benefits their organization would expect from being more data-driven, more than 56% of respondents indicated the option of improving business agility. Improving existing or developing new products and services was also a top acknowledged benefit.

However, generating insights from data more quickly does not happen overnight. Workers that use data and content in their daily roles especially depend on the ongoing management and curation of those data sources. Based on the same survey as above, data analysts currently spend 54% of their work hours simply finding and preparing data for analysis. Other roles, such as data scientists, also struggle in this regard. Time spent finding and preparing poorly managed data is a matter of wasted talent and money in many cases, slowing the organization down.

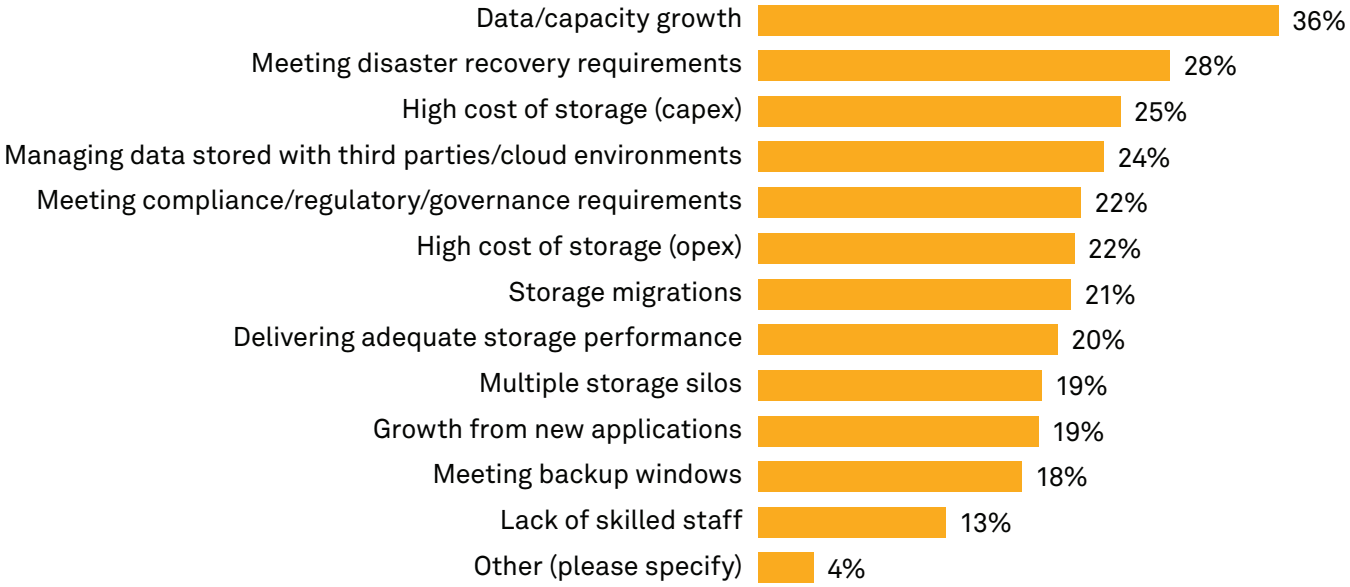
Data Storage as a Foundation of Data Leverage

Well-managed data storage is ultimately a foundational pillar for the successful downstream leverage of data in enterprise insight initiatives, potentially affecting nearly every worker either directly or indirectly. Storage is an example of a critical component warranting optimization because it doesn't just serve as a method for backup in the case of a disaster or failure. It also serves as the logical hub for the organization's informational assets – assets that are growing more important to every worker and role within the organization.

Particularly in the case of unstructured content, storage practices and technology actively support the productivity of workers across the organization and all lines of business. Problems in data storage, then, have a domino effect on business agility and organizational decision-making. Resultant scenarios can include hampered productivity and even poor-quality data for analysis.

Today, there are still numerous functional challenges in effective data storage. Cost and expenses don't necessarily tell the full story of storage pain points and inefficiency. Specifically, as storage costs have generally come down over time and data has shifted to the cloud, other key problems have emerged.

Figure 2: Top Storage Pain Points



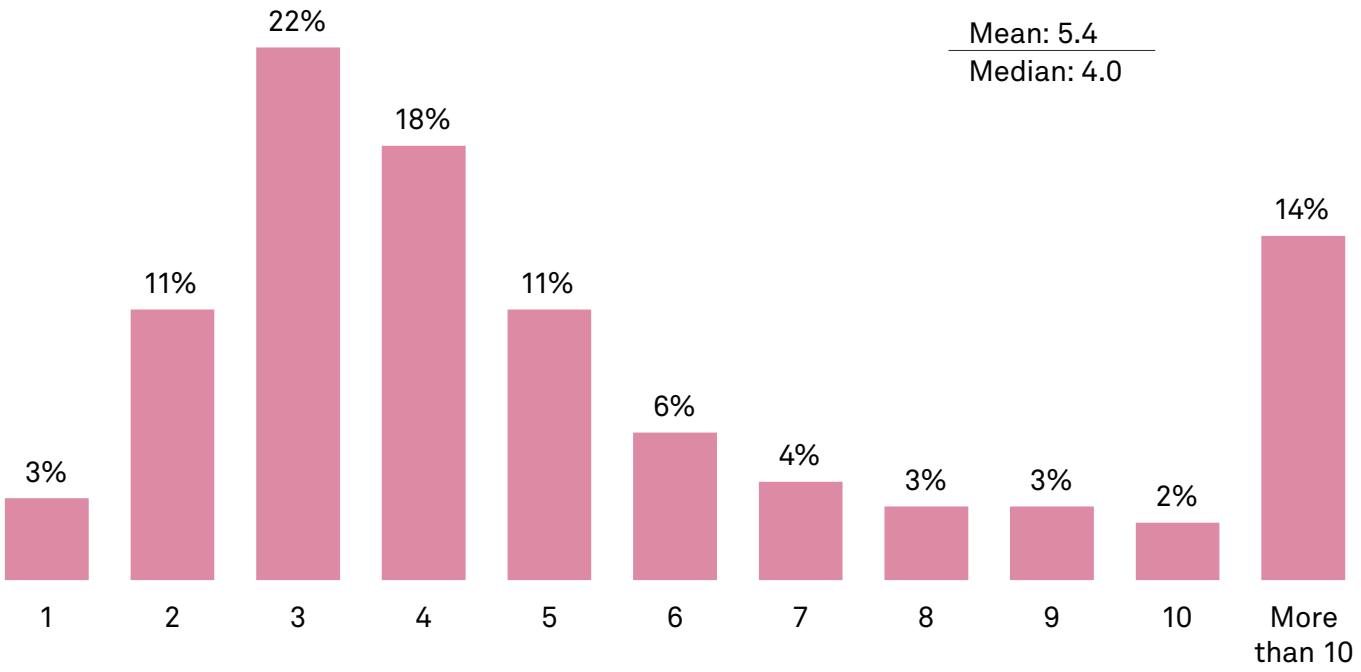
Q. What are your organization's top pain points from a storage perspective? Please select top three choices that apply.
Base: All respondents (n=444)
Source: 451 Research's Voice of the Enterprise: Storage, Data Management & Disaster Recovery 2021

The sheer growth in modern data volumes has outpaced most perceived benefits in storage cost savings. So while storge costs continue to fall, data volumes themselves continue to rise at a quicker rate in nearly all cases. Data capacity/growth, specifically, is the top cited pain point in storage, at 36% of respondents. High cost of storage (capex) still comes in at third place; the two pain points are highly interdependent.

Complexity is another factor, as indicated by several of the perceived storage pain points today. IT environments are simply more distributed and diverse than they have ever been, and this is persistently associated with numerous challenges and areas of recurring friction. For example, disaster recovery can be more difficult with more data sources and more data repositories spread across cloud and hybrid environments. Managing data stored with third parties and/or cloud environments is fraught with legal and compliance challenges, as well as nuances in responsibility. And meeting those specific regulatory, compliance and governance requirements often becomes more difficult in the absence of a monolithic, centralized architecture. Distributed architecture means different repositories and silos often have very different mechanisms for control of data, causing mismatches in policy and execution.

Increasing data volume and architectural complexity also forms a challenging feedback loop of sorts, creating yet more copies of data and further undermining the ongoing efforts of consistent data management. Most organizations today have multiple copies of business data scattered across primary storage, archives, backups and other locations both on-premises and in the cloud.

Figure 3: Copies of Business Data Across the Organization



Q. How many copies of your business data exist across your organization, both on-premises and in the cloud (e.g., primary storage, backups, archives, etc.)?

Base: All respondents, limited sample (n=291)

Source: 451 Research's Voice of the Enterprise: Storage, Data Management & Disaster Recovery 2021

The average number of reported copies of a single piece of business data, within organizations, is 5.4, and more than 30% of respondents reported that their organization has more than five copies of individual pieces of business data. And it's likely that these are underestimates. Often, extra data copies are hidden in SaaS applications, repositories and even in local storage on endpoint devices that are difficult to accurately account for in real time. These multiple copies of data can spawn confusion over data provenance/lineage, create problems with data quality, slow down productivity, increase regulatory risk and drive up storage costs over time.

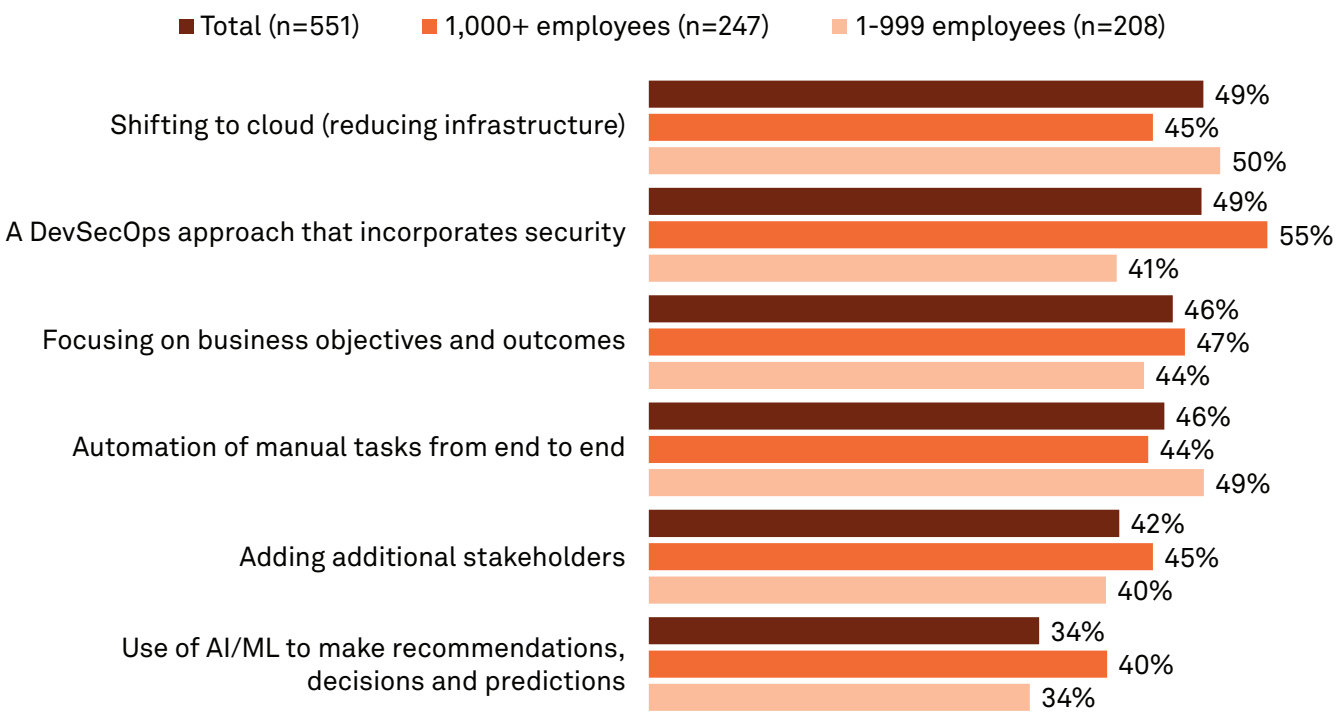
Given these considerations, data copy management and effective storage techniques are not just nice to have; they affect nearly all aspects of the way an organization interprets its informational resources. Tackling some of the persistent, underlying challenges facing data storage is an important step in adapting the business to ultimately become more data-driven.

The Need for Alignment of Business Objectives

It stands that with near-unanimous agreement that data is becoming more important to organizational decision-making, there theoretically should also be stakeholder alignment in objectives to become more data-driven. As many of the challenges and variables discussed so far would suggest, the challenge isn't in agreeing to become more data-driven; the challenge is in the process of HOW to become more data-driven. Existing architectural investments, entrenched business practices, tribal preference for specific tools and products, and ideological divisions between those who plan and those who execute can all ultimately undermine alignment.

Looking at both DevOps and storage functions provides a useful illustration of this apparent ideological division within organizations.

Figure 4: Top Priorities When Refining DevOps Strategy



Q. As your organization refines, improves and expands its DevOps strategy, which of the following are top priorities? Please select all that apply.
Base: All respondents
Source: 451 Research's Voice of the Enterprise: DevOps, Organizational Dynamics 2020

Since the advent – and widespread adoption – of agile methodology in development, the use of the cloud has become central to development practices within organizations. When DevOps-involved survey participants were asked about top priorities, shifting to the cloud (reducing infrastructure) was the top overall priority when refining DevOps strategy for organizations. Across businesses of all sizes, nearly half (49%) reported this priority.

However, this sits in stark contrast to the reality of current storage practices. Based on 451 Research's Voice of the Enterprise: Storage, Data Management & Disaster Recovery 2021 survey, an average of only 37% of an organization's data is being archived on public cloud infrastructure today. Nearly one in five (19%) storage respondents reported that their organization archives none of its data on public cloud architecture. Clearly, there is a degree of mismatch between priorities and practices across business functions, a mismatch that needs to be remedied if organizations expect to be able to leverage data more effectively.

Gaining stakeholder alignment across the organization with regard to data management and data leverage

requires making sure all relevant voices are heard. Existing technology investments are sometimes tied to external influences and requirements rather than just ideological preferences, and understanding these external factors is important for planning ahead and strategically moving forward. Consider: what are the regulatory compliance requirements? Do contractual agreements or business partnerships influence needs?

Focusing on areas of agreement in data management is a good way to lay a baseline strategy for areas to concentrate resources. Cloud compatibility of data management and analytics is a good example of this agreement. Based on 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2020 survey, 72% agree that the ability to run data management/analytics software on multiple cloud/datacenter environments is an important consideration for their organization when selecting a new data management/analytics vendor. Flexibility of options is almost always beneficial to the organization and can accommodate shifts in strategy over time.

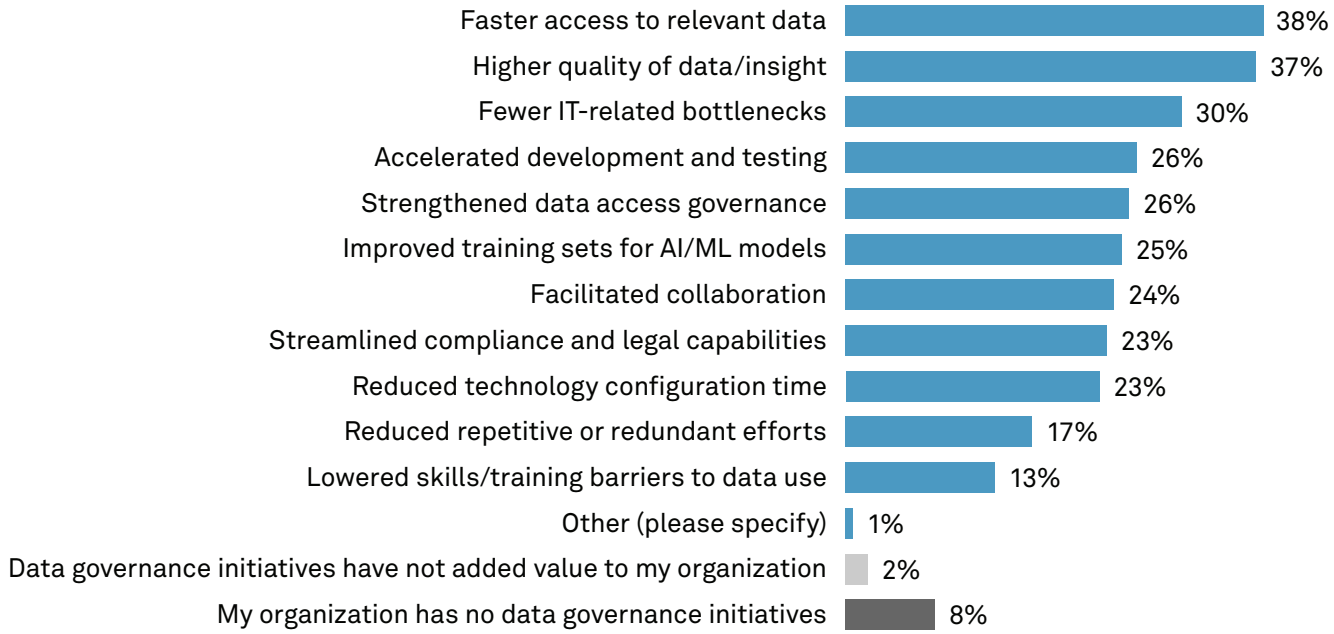
What many enterprise IT desires ultimately boil down to is to have better control mechanisms for data. Better control of data ultimately benefits all data stakeholders – stakeholders that are growing in number each day. There is a need to focus on more granular governance capabilities that can operate within modern IT architectures.

Delivering Business Value from Better Governance

Solving the data governance and data control challenge rarely can be done with technology alone, at least in isolation. The building blocks of the data-driven enterprise include people, processes and technology; all interdependent elements need to be equally supported and accounted for. Once the organization achieves stakeholder – people – alignment, it can start to focus more on the processes and technology that help achieve the consistent availability and value of organizational data.

Today, data governance is one such orchestration effort that is paying dividends for organizations, incorporating the alignment of people, as well as the support of appropriate processes and technologies. Historically, data governance was often thought of as a function inherently tied to compliance or legal requirements – at best a nuisance, and at worst an active barrier to data-derived insight. However, this perception is no longer the case for most businesses. Organizations are realizing that ongoing processes and supporting technology associated with data governance actually serve as catalysts for data-driven outcomes.

Figure 5: Business Value Delivered by Data Governance Initiatives



Q. How have data governance initiatives added value to your organization, if at all? Please select all that apply.
Base: Data management respondents, abbreviated fielding (n-149)
Source: 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2021

Key reported benefits of data governance initiatives today include faster access to relevant data, higher quality of data/insight and fewer IT-related bottlenecks, a far cry from the reactive compliance motives of yesteryear. In fact, these overall top three responses remain consistent from when the same question was asked in 2020. As the net total of workers accessing and using data within organizations continues to increase, these long-term benefits are critical to ensuring success.

Good data governance can, in effect, create a flywheel effect that benefits the entire business. Good data governance helps users more easily access and leverage good-quality data, which cuts down on inefficient human practices such as creating excess copies of business data, which in turn makes data governance efforts easier to sustain long-term – a virtuous, self-supporting cycle. Reduced redundancy in human effort, particularly with unstructured content, is another key benefit for organizations.

For the modern organization, data governance provides a common fabric for both more traditional 'reactive' requirements and more 'proactive' value-added initiatives. More reactive needs such as compliance share the same underlying requirements for enterprise data control as more proactive needs such as self-service enablement and data science programs. Without granular control of data, it is difficult to meet a requirement for compliance reporting or visibility, but it's also difficult to achieve the consistency and repeatability associated with a successful analytics or insight initiative.

Examples of shared objectives are common between 'reactive' and 'proactive' functions. Auditability for compliance, for example, requires the same underpinning technical mechanisms as data lineage and provenance capabilities that are useful for impact analysis in analytics programs. Lifecycle management for data minimization purposes can help pinpoint 'dark data' and improve data quality for ongoing efforts related to machine learning model training. Automated data access restrictions to meet privacy requirements, furthermore, can actually facilitate self-service visualization programs by allowing individuals to see and access what is immediately relevant and appropriate to them, without the bottleneck of IT.

Data governance technology, to serve all these purposes, needs to be architecturally flexible and

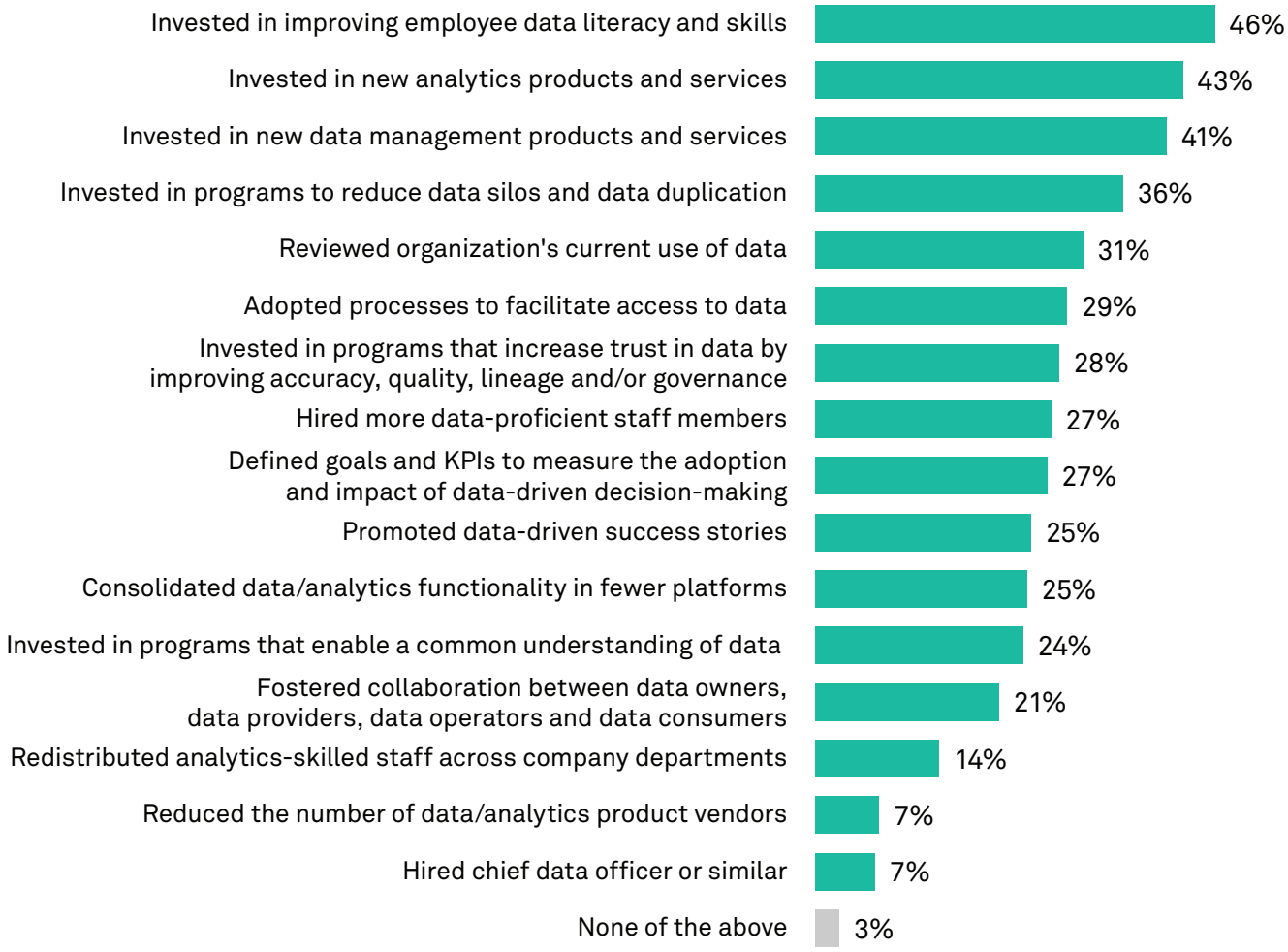
accommodating of multiple end-user personas. Focus on diverse data types, multiple data storage sources, cloud compatibility and ease of use are all key objectives. Technology alone is rarely the solution, at least in isolation. However, it can be a critical component of a successful strategy. Therefore, strategic selection criteria will consider not only baseline technical requirements, but also how that technology fits into – and supports – the people and existing processes of the organization.

Data Democratization and Building a Data Culture

The concept of organizational data culture is becoming more defined – the ongoing optimization of interactions between those people, processes and technology such that all relevant workers are empowered to drive business value via the access and leverage of data. Data awareness, data literacy, data sharing and data collaboration are all key characteristics of a thriving enterprise data culture.

Organizations today are making clear efforts to support and expand their data cultures, sometimes in slightly unexpected ways.

Figure 6: Steps Organizations Have Taken To Accelerate Data-Driven Decision-Making



Q. What steps has your organization taken to accelerate data-driven decision-making? Please select all that apply.
Base: Analytics respondents (n=216)
Source: 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2021

Data culture, in some cases, is incorrectly assumed to simply mean the pervasive use of visualization software by line-of-business users. And while analytics products are indeed an important piece of the puzzle, the effort to build a data culture requires many more input variables. Today, when enterprise survey respondents are asked to report the steps their organization has taken to accelerate data-driven decision-making, top responses also include 'investment in new *data management* products and services' rather than simply analytics. This speaks to the foundational nature of ongoing data management and governance practices and technology in a successful data culture. These efforts remain consistent. The top three responses in this survey repeat the top three responses in a 2020 survey cycle, when participants were asked to indicate what steps had been made to improve organizational data culture. While the order of the top three responses shifted slightly, the rankings remain tightly clustered.

Creating structure for data and content sources is critical. Examples of data management investments include technology such as data catalogs, data classification tools and mechanisms for the secure access and sharing of data resources. Data, and content, needs to be firmly controlled in order for workers to be able to quickly and safely leverage the information that is most appropriate and relevant to them. The control that helps secure data is also the control that allows workers to easily navigate the data ecosystem and find what they need.

Other key considerations for data culture include employee data literacy, as well as efforts to reduce data silos and data duplication – all things that immediately contribute to worker productivity with data. Results from past surveys suggest line-of-business decision-makers are often the primary focus for investment in these new products and services, and this is telling. Line-of-business groups are the largest population of workers within the organization, and today, much of the collaboration and sharing of data and content occurs with them, not necessarily IT. Supporting technology must have their needs in mind.

Conclusions

The ongoing integrity of business insight depends on the integrity of the ongoing management of data and content within an organization. This is not just an IT problem, nor simply a technological one. Workers today, in a functioning data culture, depend on the accessibility and reliability of data and content as they seek to be productive in their roles and – increasingly – contribute to business value via data-derived insight. A high-level approach will consider not only supporting technology, such as cloud architecture and storage considerations, but also simply the way people prefer to work, including favored processes and tools.

The desire to increase productivity with data should be a commonly held value shared across the organization, regardless of business unit or function. The effective management of content is clearly an immediate facilitator for productivity with workers, but broader objectives should seek to gradually apply more cohesive mechanisms to manage all various data types as a common resource, regardless of scale. Ongoing data governance efforts are an obvious way to deliver value today, in ensuring data reliability as well as risk mitigation for a business ever-more dependent on data, but communication of objectives and stakeholder alignment are also critical.

Technology provides necessary support but cannot ameliorate existing ideological divisions within the walls of an organization. Leadership buy-in and guidance in achieving enterprise-wide leverage of data is necessary so that appropriate supporting technology can be selected that is mutually beneficial to all data stakeholders as the number of data stakeholders continues to grow.

CONTACTS

The Americas
+1 877 863 1306
market.intelligence@spglobal.com

Europe, Middle East & Africa
+44 20 7176 1234
market.intelligence@spglobal.com

Asia-Pacific
+852 2533 3565
market.intelligence@spglobal.com

www.spglobal.com/marketintelligence