

TRENDS IN DATA MANAGEMENT

February 2020

Modern technology strategies have become a balancing act. On one side, there are emerging trends and techniques that can give businesses an edge in the digital marketplace. On the other side, there are the foundational technologies and practices that don't generate as much buzz but are still critical for digital transformation. The management of corporate data is a prime example of the foundational side. Digital data has obviously been present for decades, but proper management has become critical in an age of artificial intelligence, predictive analytics, and customer personalization. This study explores current trends in data management and future steps that companies plan to take in order to deal with an expanding set of data issues.

KEY POINTS

Businesses are in the early stages of building their data management strategies

Only 25% of companies feel like they are exactly where they want to be with their corporate data management, down from 31% in 2015. Although digital data has long been a part of IT operations, there has not been much focus in terms of job roles or defined components. The hype around big data highlighted a need for comprehensive data policies, and businesses are now beginning to build discipline around capturing, processing and analyzing data.

There are many challenges in tying data management to business strategy

The list of challenges that businesses are facing in building a data strategy shows how important it is to have an established process. For example, 38% of companies cite a desire to speed up their data analysis, which involves both infrastructure and process. Similarly, 37% cite the need to integrate data throughout the organization, which involves having the right storage solutions but also the right workflow. Across infrastructure, skills and policies, there are a range of issues that businesses have to solve to properly manage their data.

New and improved skills are needed

Only 44% of companies say that they have internal IT employees who are dedicated to data management or data analysis. While there has been a focus on newer job titles like data scientist, there is also opportunity around more traditional roles like database administrator. Some firms may need to build these roles for the first time, while other firms with these roles in place may need to expand the skill set. Since building all the necessary skills internally is a costly endeavor, partnering will also be a viable option, with 65% of companies expecting to explore third party data services for the first time.

Blockchain and other DLT solutions provide a promising option for data structures.

Spurred by the use of blockchain in cryptocurrency applications, companies are exploring different solutions that use blockchain or other distributed ledger technologies as the underlying foundation. Adoption numbers are currently skewed by lack of familiarity and marketing hype, but applications such as digital identity or smart contracts could lead the way in demonstrating the benefits of DLT.

MARKET OVERVIEW

Much has been made of the importance of data in the digital economy. Data has been called the “new currency” and the “new oil” of business—an asset so critical that it undergirds the entire operational structure. Yet for all the value being ascribed to digital data, there is less consideration for how that data is managed.

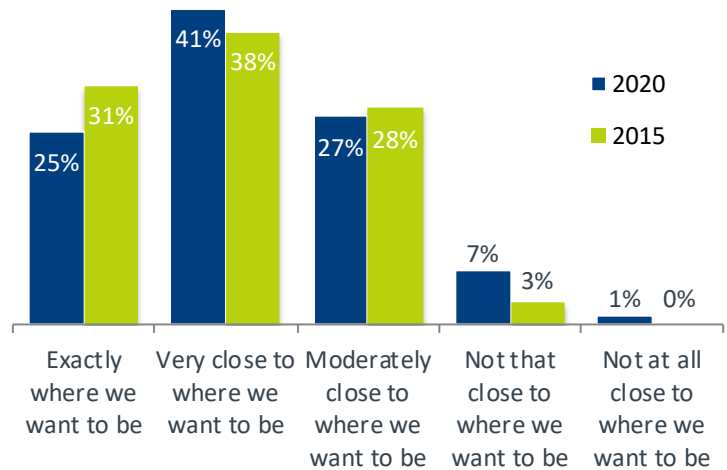
There are several possible reasons that comprehensive data management has not received the attention it deserves. First and foremost is that the view of data as a critical asset is a relatively recent phenomenon. Obviously companies have been collecting and analyzing data for decades, but they did not necessarily treat this analysis as a primary discipline. Industry observers believe that the first Chief Data Officer (CDO) was appointed in 2002, and a 2018 study from NewVantage Partners showed that the number of firms employing a CDO had reached 62.5% after hovering at a mere 12% in 2012. The role of data scientist is one of the hottest jobs in technology, but it is also a recent creation, born during the big data craze. Earlier, there was not an equivalent “data engineer” role in the same way that companies employed network engineers or server administrators.

Another reason that data management gets less attention is that it has many different facets which are major topics in their own right. Privacy is clearly something that most businesses are focused on today as they determine the best way to be transparent while still maintaining a competitive edge. Data breaches are top of mind as companies try to ensure ongoing operations and continued trust from their customers. Each of these areas has been the topic of much debate and dedicated regulation (such as GDPR or the California Consumer Privacy Act), but they are still just pieces of an overall data management strategy.

Finally, data management specialists in the past typically came from development groups. In the interviews conducted for CompTIA’s *Functional IT Framework*, many CIOs reported that their lead database administrator had previously been a software developer who built responsibilities around data. Since software development was not a core competency for many companies in the age of packaged software, it makes sense that the related field of data management is new territory.

Whatever the reasons, data management has not been a corporate function in the same way as a topic such as cybersecurity. That may be changing, though. Along with the rising number of CDOs, the strong desire to use emerging technologies like internet of things and artificial intelligence points to a need for holistic data management. These trends require consistent data as inputs and generate new forms of data as output. Without plans and policies around data handling, it is impossible to leverage these cutting-edge technologies. This is why CompTIA’s IT framework ultimately included data as a primary IT pillar along with infrastructure, software development, and cybersecurity.

Views on current data capabilities



Given the lack of focus on data management, it is no surprise to see that companies have not made much progress in the past five years. In fact, the number of companies that feel they are exactly where they want to be with their data capabilities has taken a step backward. As with most issues related to digital transformation, getting everyone on the same page is key to moving forward.

There is a distinct difference in viewpoint across workforce segments. The most satisfied group is the executive level, with 41% reporting they feel their company’s data capabilities are exactly where they should be. Without a top-down mandate, it can be difficult to build a corporate strategy. Satisfaction takes a big hit for the next group, IT staff. Only 23% of these employees feel that their data capabilities are ideal, likely because they are the most aware of areas for improvement. Finally comes business staff, with 17% reporting the highest level of satisfaction. This group is the most likely to understand the shortcomings of a subpar data approach, but they also have the least ability to pull data from all parts of the organization.

At first glance, it may be surprising to see that company size is not a major factor in satisfaction around data practices. The highest satisfaction is found in medium-sized companies (100-499 employees), with 28% reporting they are exactly where they would like to be. Large companies (500+ employees) are not far behind, with 26% reporting high satisfaction, and even 21% of small companies (less than 100 employees) fall into the same category. While there are different resource capabilities across these firms, there are also different datasets and data needs, and the two sides balance out.

If data is truly a critical resource, then it needs to be managed properly. If satisfaction with data handling is relatively low, then there is plenty of room for improvement. This report will examine the different facets of data management and the steps that companies can take to best utilize this valuable asset.

SETTING THE STAGE FOR DATA MANAGEMENT

With data at the heart of digital transformation activities, it makes sense to begin discussions around data management with a general look at the state of enterprise technology. Overall, satisfaction with technology tracks satisfaction with data practices: 31% of companies feel completely satisfied with their current technology, 50% are mostly satisfied, and 17% feel that their technology is simply adequate. Here, there is a little more differentiation by company size. There are more large companies that are completely satisfied (36%) than there are medium-sized (28%) or small firms (29%). While the problems are more complex, wider availability of resources and deeper pockets seem to help bridge the gap.

Diving deeper on both sides of the satisfaction spectrum, it becomes clear that data is not seen as a driving factor for either positive or negative feelings. Having well-defined data management practices falls in the middle of the pack as far as reasons that employees feel satisfied with their company's technology, with 48% citing this as a key contributor. It is worth noting that the top drivers for positive sentiment center around the effectiveness of existing systems.

On the other side, poor data management practices are one of the least common reasons for people to feel dissatisfied with their current technology approach. Only 18% of respondents cited this as a significant factor. Again, the effectiveness of existing systems was a top driver for negative feelings.

To be sure, having existing systems that meet the needs of the business is critical. Without functional systems and solid maintenance processes, pushing the envelope on technology is an exercise in futility. However, the transformation into a digital organization requires a greater appreciation for the way that the entire architecture comes together. Cybersecurity is one part of that, as the protective layer surrounding the strategy, and there is an emphasis on security as a corporate priority. Data is another part, as the connective layer between applications and the fuel for making intelligent decisions. As discussed in the market overview, data management has not yet received the same level of emphasis.

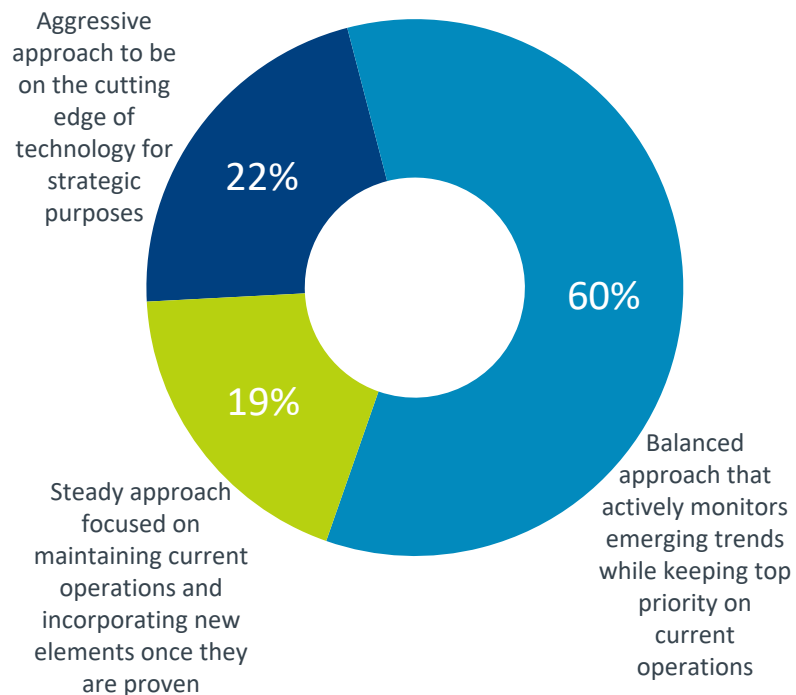
As businesses embrace digital operations and build data practices, there are four areas they must focus on. Those companies that have installed a CDO have likely assigned these responsibilities to that position, but all companies can discuss these areas as part of a data strategy regardless of who is executing that strategy.

- 1. Data sourcing.** With data coming from a wide variety of sources, the first step is to understand all the inputs. Part of this process involves ensuring the data is clean and high quality, especially with the growing risk of data poisoning as a form of cyberattack.
- 2. Data processing.** From data consolidation to storage solutions to processing schedules, this step is where

infrastructure must be matched with data flow. There are myriad tools available for different types of data and different types of processing, and picking and integrating the right set is a sizable effort.

- 3. Data analytics.** This step involves analyzing the data, creating visualizations, and presenting the findings to business leaders. There is a technical component, but there is also a strong need for business acumen in order to tie the data to corporate objectives.
- 4. Data security.** Across the whole process, the data function must work closely with the security function to ensure data integrity and minimize the risks of data loss or breaches.

Organizational approach to technology



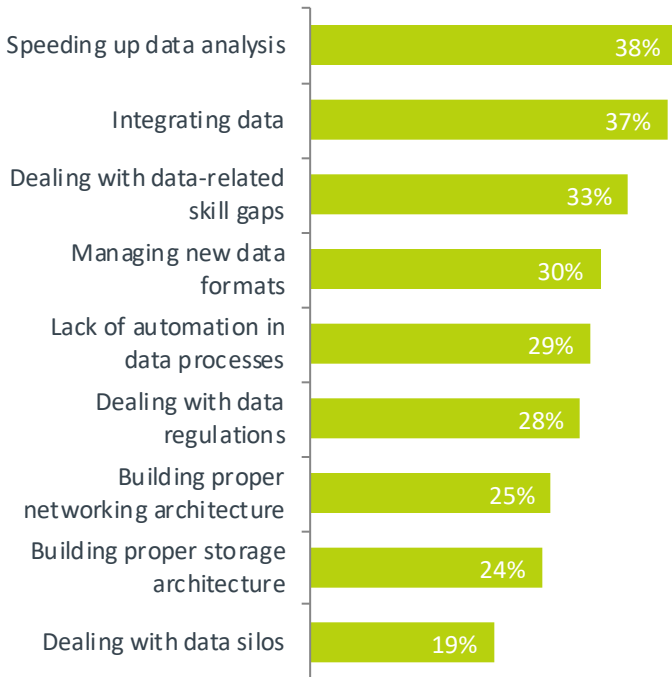
While these steps can all be applied to current IT operations, it is critical to establish a process before moving forward with emerging technology initiatives. Most companies are paying some attention to new trends, and a solid data strategy can help bridge the gap as firms find the balance between maintenance and innovation.

Furthermore, most emerging trends feature data techniques that are slightly more advanced. Whether it is turning IoT data into an automation scheme or accumulating the appropriate data for AI algorithms, new technology does not function as well if the data is not being handled properly.

DATA MANAGEMENT CHALLENGES

For most companies, there is not much to build on when it comes to establishing comprehensive data practices. On one hand, this creates a blank slate where businesses can create procedures without the burden of legacy thinking. On the other hand, everything has to be built from scratch, introducing a full range of problems in every part of the process.

Challenges in building data management practices



The top challenge that companies cite is trying to speed up data analysis. In a digital environment that has become characterized by rapid changes, there is a growing demand for instant results. However, there are several bottlenecks that can impact turnaround time. Even as new tools were coming online during the beginning of the big data hype, one of the major hurdles was producing real-time analysis. 5G networks are highly anticipated, but network speed may not be the slowest part of the path today. Getting faster results from data will require careful consideration of each piece of data architecture, along with workflow that is tightly managed.

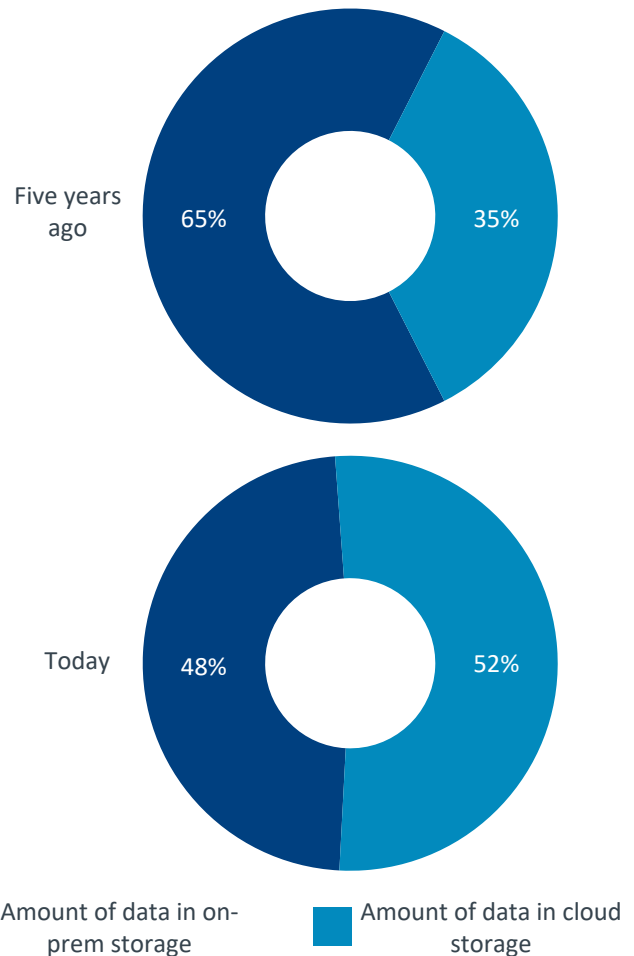
The next challenge on the list highlights the difficulty in bringing together different technology initiatives across the organization. For several years, CompTIA research has found that business units working independently on technology projects eventually leads to a challenge with integration. As a result, companies are trying to avoid pure shadow IT in favor of a collaborative approach that still gives business units some freedom while maintaining an inclusive view of all business systems. Collecting the data into a single repository is part of this, and is also critical for AI initiatives that operate on the broadest possible datasets. It is interesting to see that data silos are not widely considered a problem even though data integration is a top challenge. Given that 82% of companies

say they have a high or moderate degree of data silos, there is a clear disconnect on how problematic data silos are and how exactly they should be integrated into a common dataset.

The set of challenges in the middle of the pack addresses the scale of the data management problem. Filling skill gaps is obviously a critical need, and the topic will be examined in more detail further in this report. Handling new data formats was one of the main draws for big data tools, but it typically involves a completely separate workflow rather than an extension of any existing flow. With the scope increasing faster than skills can be added, automation must be considered as a way to deal with complexity and routine tasks. Nearly 90% of companies say they have automated some part of their data management process, from the back-end steps of analysis and reporting to the front-end step of filtering data as it comes in.

Although infrastructure issues are not listed as a top challenge, many companies are still reconfiguring their architecture to take full advantage of cloud systems. Networks are in a constant state of flux as new data sources are tied in and traffic is prioritized. Storage has clearly become a hybrid solution, with on-prem storage often used for the most critical information and cloud storage used for a wide range of operational activity and archiving.

Shift to cloud storage

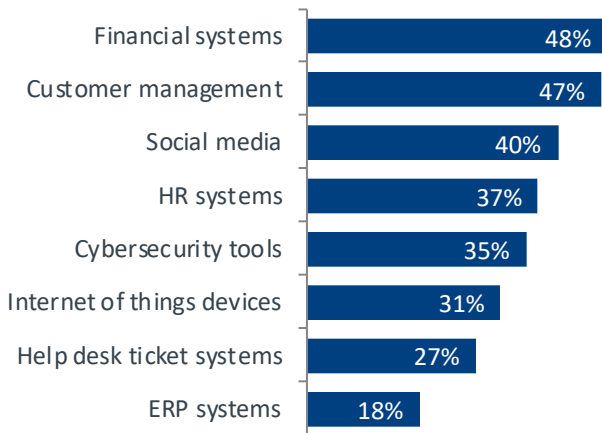


COMPONENTS OF DATA MANAGEMENT

Over the past five years, companies have slowly been building an approach to data management that incorporates all the data throughout the organization and creates formal processes for consistent operations. The big data trend has largely faded into the background as businesses realize two things. First, they must have a solid foundation of data management before taking on drastically new pieces. Second, it is better to have a single overarching approach rather than treat big data as something separate. Currently, just over half of all companies even recognize big data as part of their strategy, with 14% taking a modern approach of treating all data as one combined concern.

The first component of a data strategy is to fully understand the different sources of data. Data sources can relate to business operations (financial systems, ERP) to customer profiling (customer management, social media) to IT concerns (cybersecurity, help desk). Different departments may be finding new uses for their own datasets, but the optimal opportunity for insight and automation comes from pulling all the data together and finding nonintuitive connections. The types of applications producing data are currently slanted towards the most common business applications, but as cloud systems allow for greater application complexity, there will be a wider variety of data.

Common sources of data

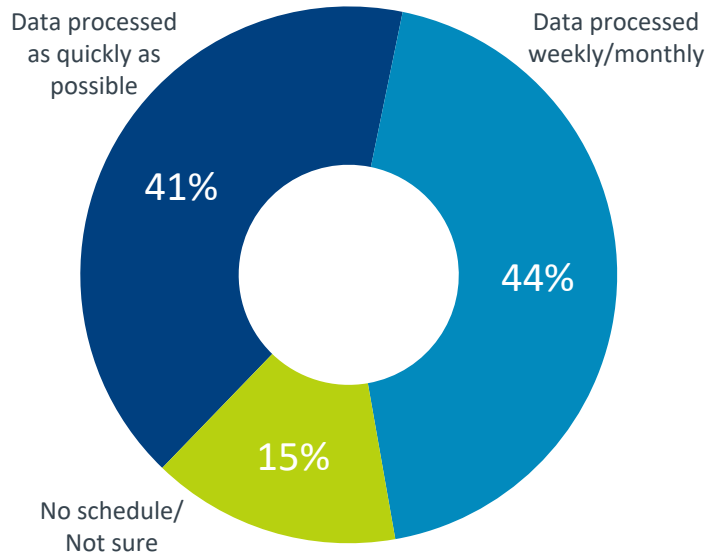


The second step of the data management flow is processing and organizing the data. This stage has expanded significantly in recent years, with companies moving beyond SQL tools and a standard Extract/Transform/Load (ETL) process into NoSQL, NewSQL, in-memory databases, and other tools that can handle unstructured data along with structured data.

One of the most important metrics for the processing stage is the speed of moving from raw data to analysis. The default goal is for real-time processing, but the devil is in the details. Just as there is increasing cost in moving from four nines to five nines in reliability, moving closer to real-time analysis demands sizable investment. Among those individuals who are aware of their company's data processing approach, there is a relatively even split between those who say they are

processing data as quickly as possible and those who currently rely more on batch jobs happening on a weekly or monthly basis.

Approach to data processing



The final step is data analysis, where the goal can be either understand trends from the past or predicting the direction of the future. Again, speed is important here along with matching analysis and visualization to the audience and the overall business flow. It is interesting to note that at this time, the heaviest concentration of stakeholders in the data management process is at the department level. Moving forward, there will be a stronger demand for data analytics at the highest levels of the organization.

Three major trends will shape the data management process over the next twelve months. First, natural language processing will open the door to data analysis being performed by non-technical staff. Without having to know a specific coding language, employees can query the data using terms that make sense to them.

The second trend is more technical. With the increased interest in relationships between disparate data points, graph databases will rise in popularity. By far, the most popular tool used for data manipulation and analysis is the spreadsheet, which has many limitations in dealing with today's data. Graph databases such as Apache Cassandra, Amazon Neptune, and Neo4j use a foundation of nodes and relationships to bring more performance, flexibility, and agility to data analysis.

Finally, artificial intelligence will bring more capability to data processing. Whether it is automation of certain steps or augmented analysis using machine learning algorithms, companies will leverage AI to handle the massive scope and complexity of their data and to assist their data management teams in driving value.

DATA MANAGEMENT TEAMS

The concept of data teams is relatively new. Given that the CISO position was created in the mid-1990s and the CDO position was created in the early 2000s, it makes sense that data teams would be lagging behind security teams. Only 44% of companies say that they have internal IT employees who are dedicated to data management or data analysis. Both large companies (52%) and medium-sized companies (51%) are leading the way in creating a dedicated function. It is actually somewhat surprising to see that 33% of small businesses also say they have dedicated data staff, considering many of these firms do not have dedicated IT staff.

Looking at the skill sets that companies have for their data needs, there is obviously a concentration of the more traditional skills. Database administration is the fundamental skill needed to manipulate data, and there are many established paths for growing a career as a database administrator (DBA). Some companies may consider their general infrastructure team as having data infrastructure skills, but there is also a need to focus specifically on data structures, especially with the wide variety of structures needed for different types of data.

The surprise in the top three is data analytics. Not only does the current number look low, but there appears to be a low number of companies planning to add this skill or improve capabilities. One explanation could be confusion over titles. Many companies may employ “business analysts” rather than

“data analysts,” even though the skills are very similar. Similarly, companies may be planning to build skills in “data science.” The exact skills needed for any job role is always a subject of debate, and well-defined certifications can help companies be sure they are getting the skills they need.

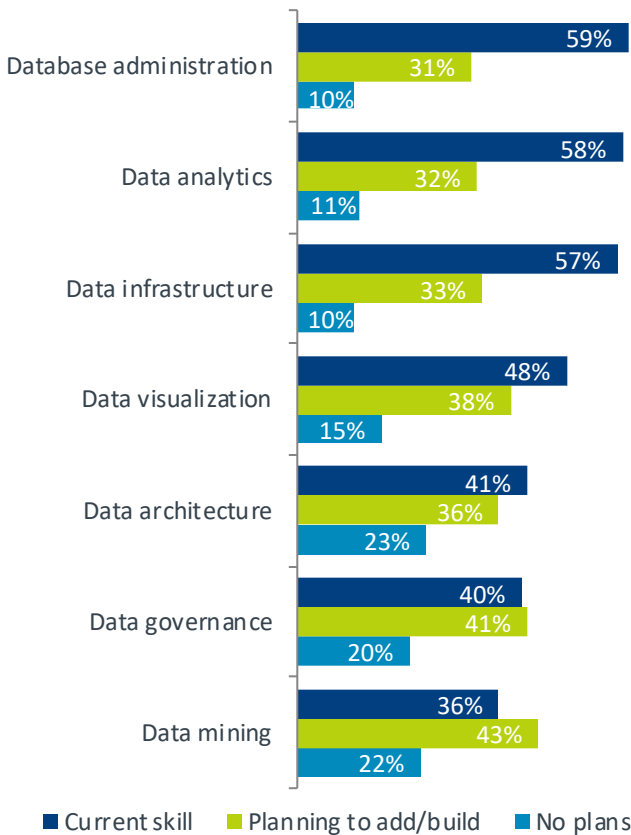
Beyond the basics, there is more granularity. Correspondingly, these skills are less common with more of a future demand. Data visualization is an interesting role since it has far less of a technical component and relies on business knowledge and artistic acumen. Data governance is quickly becoming more critical as there are more regulations around digital data and the regulations differ across state or national borders. Data mining speaks to the fact that companies are casting a wide net around all their data in hopes that they will find some unknown treasures.

As with the cybersecurity space, the wide range of skills needed for comprehensive data management is quickly growing larger than many companies are able to handle. Larger companies may have the resources to build whole teams but still may choose not to invest in the full spectrum of skills. Third party assistance, then, will be a growing part of a data strategy.

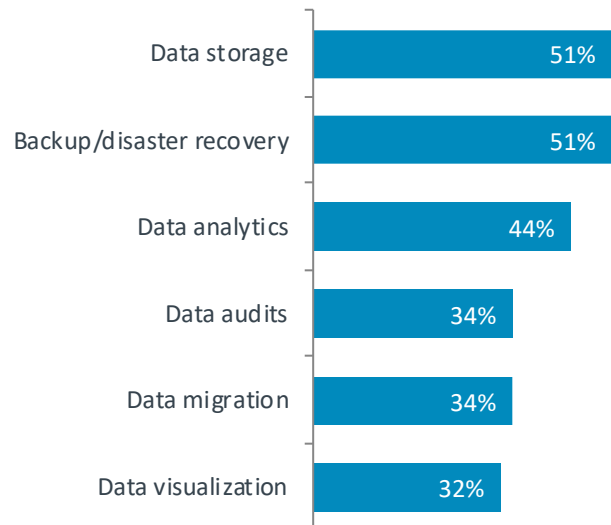
Currently, only 27% of companies are using a third party for data services. The most common tasks being performed by third parties are backup/disaster recovery and storage. Along with cloud providers, managed service providers have a good foothold in this space since they offer offsite storage. This helps mitigate risk as well as reduce the amount of routine work that a company has to perform.

Looking ahead, 65% of companies not currently using third party data services said that they were extremely likely or somewhat likely to consider a partnership in the next twelve months. For now, the expected uses are still around the strengths of storage and backup. As data management practices become more formal, other services will likely be needed.

State of data skills



Expected use of third party data services

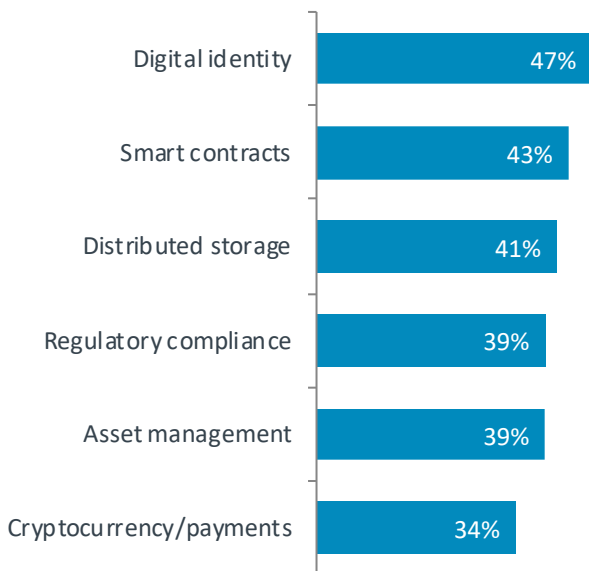


DIGITAL LEDGER TECHNOLOGY

The blockchain hype has cooled a bit, but there continues to be a high level of interest in using blockchain or other digital ledger technologies (DLT) as data structures. With the original focus on cryptocurrency, the primary question around using DLT in the enterprise has been the possible applications. However, that masks a more fundamental question around the place of DLT in the technology stack. Thinking about the function of DLT, the most logical use of this new trend is as a replacement for a database. This functionality is what gives rise to a host of possible applications, and it also leads to a host of challenges.

According to CompTIA's survey, 21% of companies say they are developing tools that use DLT as the underlying technology, and 27% say they have purchased tools that use DLT as the underlying technology. As with past CompTIA studies on emerging trends, these numbers appear high compared to anecdotal evidence about DLT adoption. The methodology CompTIA uses does not provide a level of detail to understand these company viewpoints in depth, but a lack of understanding around DLT technology combined with potentially confusing messages from vendors likely leads to overinflated estimations around the current use of DLT.

Types of DLT applications being explored

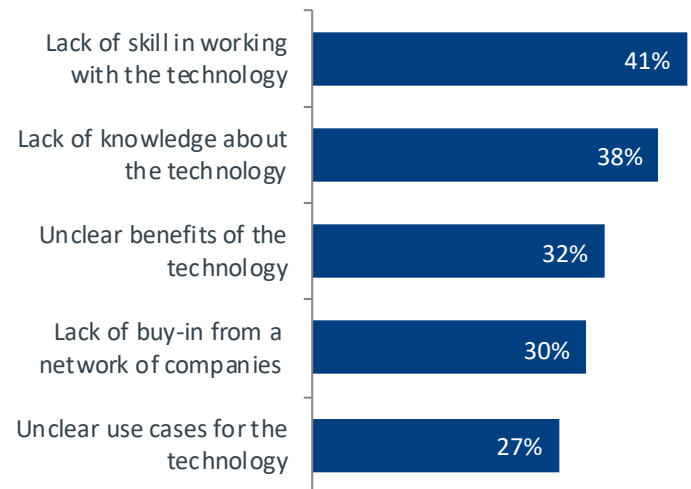


Even with inflation around adoption, the types of DLT applications that companies say they are exploring gives insight into the types of DLT applications that may eventually become widely adopted. To start, it is noteworthy that cryptocurrency ranks last on the list. Again, there is a divide between the use of DLT for cryptocurrency on the open market and the approach that companies will take. Few companies today are in the business of dealing directly with currency issues, and there is no reason that will change in the future.

The applications that are getting the most focus, then, are those applications that currently have no viable precedent in

today's technology or those applications that could be drastically changed with the introduction of DLT. There are no widely accepted solutions for digital identity, and while it remains to be seen whether the market will support universal identity solutions, DLT has more potential than existing tools. There are some current attempts at smart contracts without DLT, but DLT provides answers to questions of binding terms and highly reliable data.

Expected challenges with DLT solutions



The top challenge that companies expect to face with DLT is a lack of skill in working with the technology. This is more evidence of the main theme throughout this report of companies not understanding their overall data management. Of companies that say they are using DLT solutions, only 33% say they have taken a significant level of action in building deep internal skill or seeking external expertise. The majority have extended their existing skills. Few companies will need deep skills in DLT, just as few companies today need deep skills in designing databases. The skills needed are around manipulation and analysis of the data, whether that data is in a DLT structure or a traditional database.

One challenge that is perhaps underrated is the lack of buy-in from a participating network of companies. The true value of DLT, which takes it beyond the traditional database, is in distributing the ledger across all interested parties so that there is a consistent record of transactions. To achieve this, though, requires agreement from everyone. Obtaining this agreement is likely to be a costly exercise that reduces any benefit from eliminating a clearinghouse.

Over time, companies will learn to balance these challenges against the real-world benefits they get from DLT solutions. Today, businesses are pursuing DLT primarily for security and better data integrity. In the long run, other benefits such as transaction transparency and cost savings may become more prominent. There are clearly hurdles in the way of adoption, but eventually DLT will likely become a critical enabling technology, even if the underlying details are hidden for most applications. Properly implementing DLT where appropriate will become a key part of data management as organizations build their corporate strategies.

RESEARCH METHODOLOGY

This quantitative study consisted of an online survey fielded to workforce professionals during December 2019. A total of 400 businesses based in the United States participated in the survey, yielding an overall margin of sampling error proxy at 95% confidence of +/- 5.0 percentage points. Sampling error is larger for subgroups of the data.

As with any survey, sampling error is only one source of possible error. While non-sampling error cannot be accurately calculated, precautionary steps were taken in all phases of the survey design, collection and processing of the data to minimize its influence.

CompTIA is responsible for all content and analysis. Any questions regarding the study should be directed to CompTIA Research / Market Intelligence staff at research@comptia.org. CompTIA is a member of the market research industry's Insights Association and adheres to its internationally respected code of research standards and ethics.

ABOUT COMPTIA

The Computing Technology Industry Association (CompTIA) is a non-profit trade association serving as the voice of the information technology industry.

With approximately 2,000 member companies, 3,000 academic and training partners, 100,000-plus registered users and more than two million IT certifications issued, CompTIA is dedicated to advancing industry growth through educational programs, market research, networking events, professional certifications and public policy advocacy.



OTHER RESOURCES

RESEARCH

CompTIA publishes 20+ studies per year, adding to an archive of more than 100 research reports, briefs, case studies, ecosystems, and more. Much of this content includes workforce analyses, providing insights on jobs, skills, hiring practices, and professional development.

[CompTIA Research Library](#)



CERTIFICATION | LEARNING

CompTIA is the leading provider of vendor-neutral skills certifications and education of the world's IT workforce. CompTIA has four certification categories that test different knowledge standards, from entry-level to expert, in cloud computing, mobility, Linux, networking, security, help desk and technical support, servers, project management and other mission-critical technologies.

[CompTIA Certification and Resources](#)



COMMUNITIES | COUNCILS

CompTIA member communities and councils are forums for sharing best practices, collaborative problem solving, and mentoring. Discussions frequently revolve around the types of emerging trends covered in this report.

[CompTIA Communities](#)



ADVOCACY

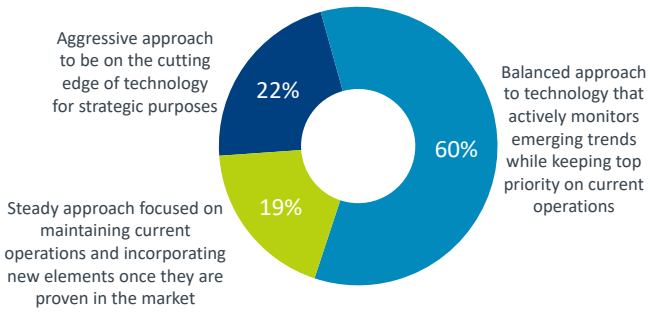
Through its public advocacy efforts, CompTIA champions member-driven business and IT priorities that impact the continuum of information technology companies – from small IT service providers and software developers to large equipment manufacturers and communications service providers. CompTIA gives eyes, ears and a voice to technology companies.

[CompTIA Advocacy](#)



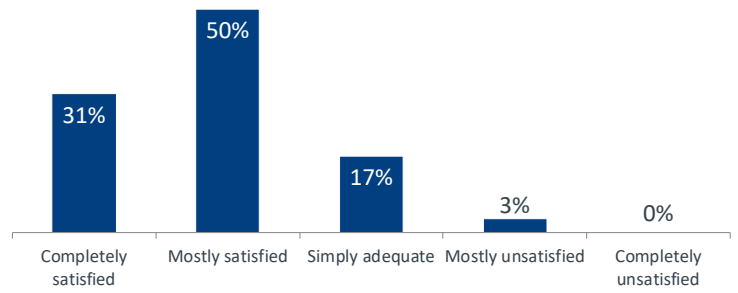
APPENDIX

Approach to Technology



Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Satisfaction with Company Technology



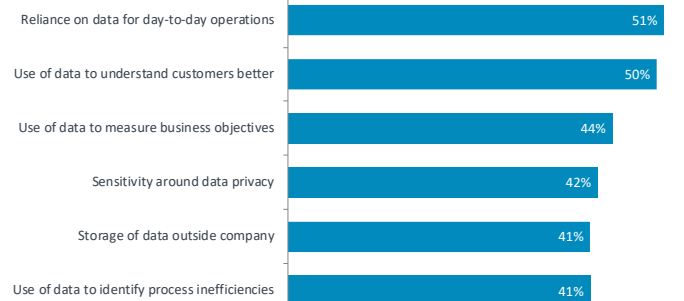
Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Reasons for satisfaction with technology



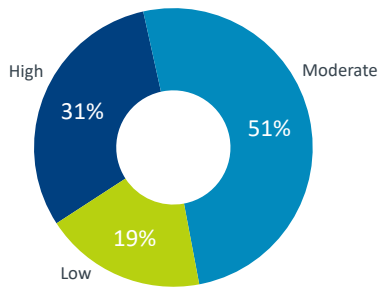
Source: CompTIA's Trends in Data Management | n = 322 IT and business professionals in the U.S. satisfied with corporate technology

Reasons for increased importance of data



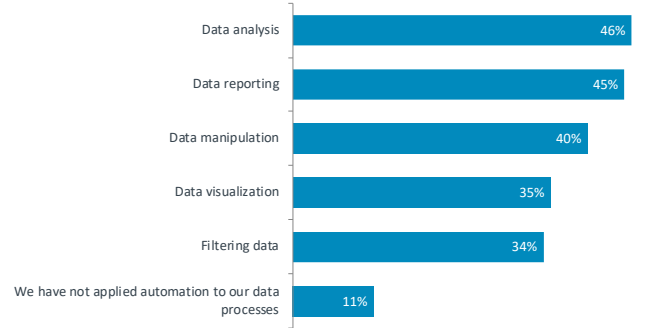
Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Degree of data silos



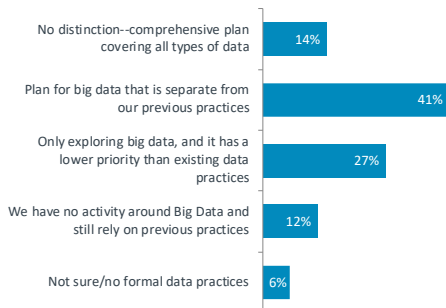
Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Parts of data process that are automated



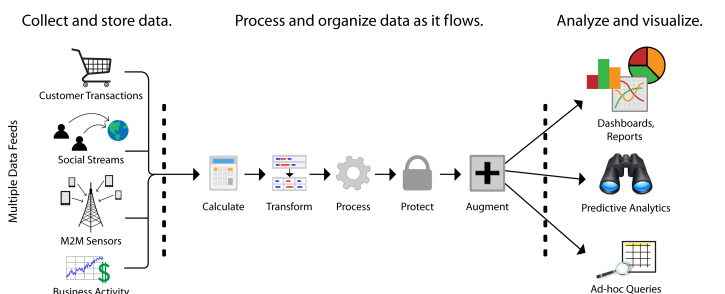
Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Approach to "big data" vs. "normal data"



Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

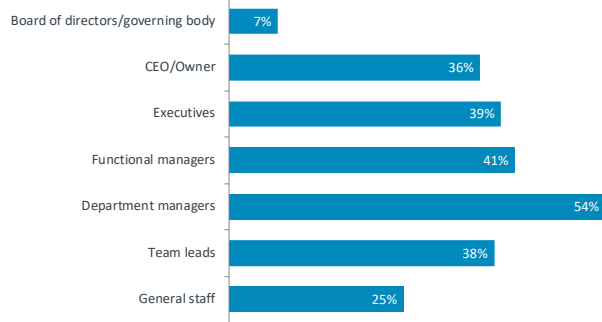
Three stages of data management



Source: CompTIA

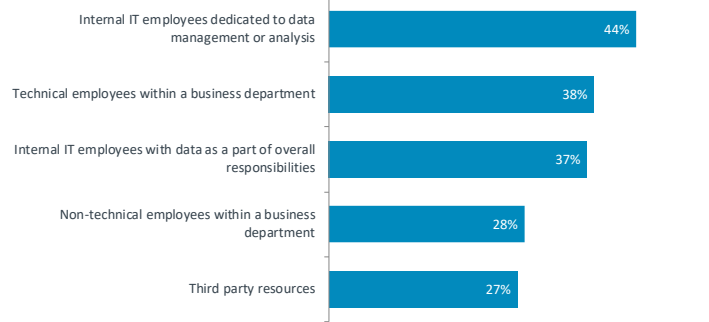
APPENDIX

Data stakeholders/decision makers



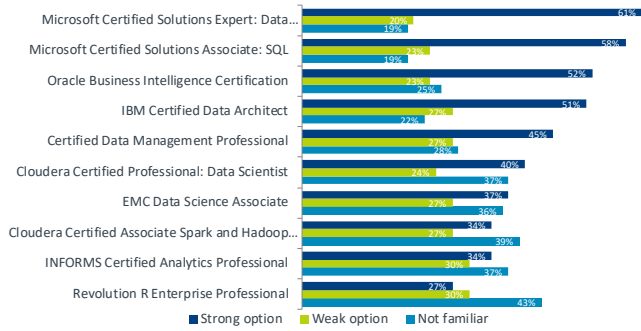
Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Groups involved in data operations



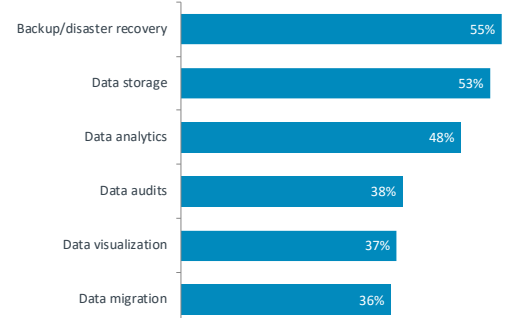
Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Views on data certifications



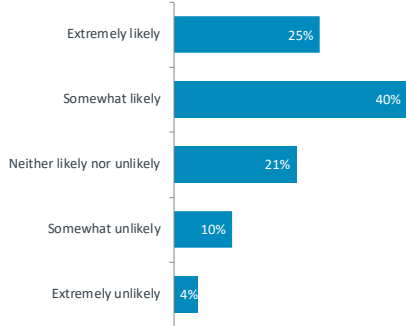
Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Use of third party firms for data



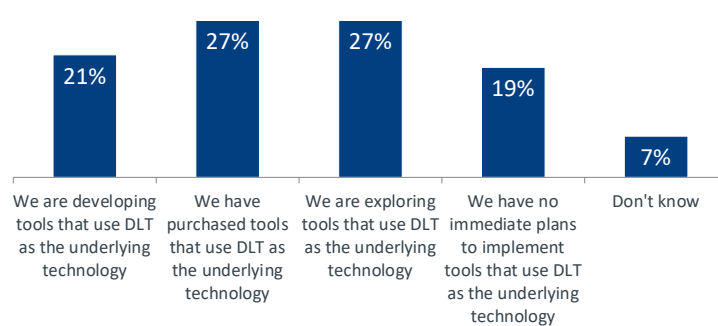
Source: CompTIA's Trends in Data Management | n = 306 IT and business professionals in the U.S. using third party data services

Likelihood of using third party firms for data



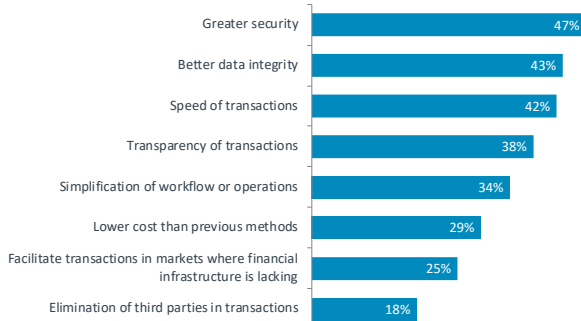
Source: CompTIA's Trends in Data Management | n = 294 IT and business professionals in the U.S. not using third party data services

Adoption of DLT



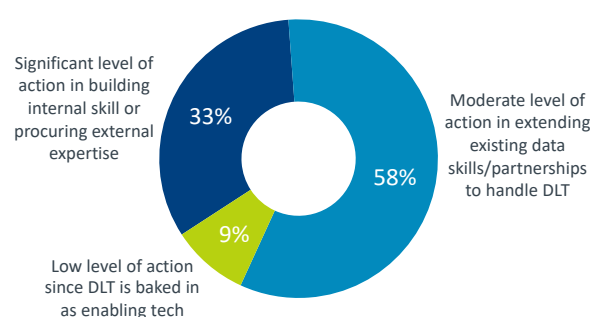
Source: CompTIA's Trends in Data Management | n = 400 IT and business professionals in the U.S.

Drivers for DLT



Source: CompTIA's Trends in Data Management | n = 297 IT and business professionals in the U.S. using DLT tools

Approach to DLT skills



Source: CompTIA's Trends in Data Management | n = 297 IT and business professionals in the U.S. using DLT tools