



The Current State of Modern Data Management: Building an Intelligent Data Enterprise to Improve Business Outcomes

RESEARCH BY:



Stewart Bond

Research Director, Data Integration
and Data Intelligence Software, IDC



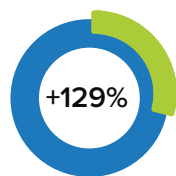
Table of Contents

Click on titles or page numbers to navigate to each section.

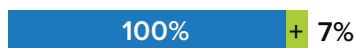
Key Findings	3
IDC Opinion	5
In This White Paper	6
Situation Overview	7
Organizational Structure and Priorities	7
Current State of Data Management Technology	8
Metadata	9
Automation	11
Measurements of Success	12
Resolving Complexity Through Maturity	14
Why? Optimized Data Management Organizations Are Measuring Better Outcomes	15
What? Optimized Data Management Organizations Have Standardized Data Management Functions and Embraced Automation	17
How? Optimized Organizations Are Setting Many Objectives and Efficiently Manage and Share Data, with Centralized Organizational Structures That Can Drive Business Value	20
Future Outlook	24
Technology	24
Automation	25
Organization	25
Challenges/Opportunities/Recommendations	26
Conclusion	28
Appendix	29
Survey Demographics	29
Building the Maturity Scale	30
About the Analyst	31

Key Findings

Organizations with the most mature data management capabilities have experienced more than 2.5 times (+129%) the improvement in data metrics (e.g. quality scores, data issue frequencies, time to delivery, data duplication, security, privacy and compliance risk) and a **107%** improvement in business metrics compared with the least mature



Improvement in data metrics

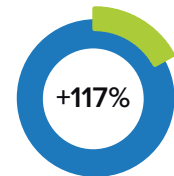


Improvement in business metrics

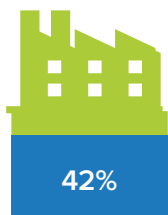
Organizations at the top of the maturity scale are measuring 100% more metrics and seeing even more improvements in data management metrics — at a factor of 2.2 times, or 117%.



Maturity scale metrics



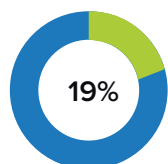
Improvement in data management metrics



Forty-two percent of organizations do not believe that all or most of the data they have is available to be used by application developers, data stewards, data engineers, data analysts, data scientists, and business analysts.



Forty percent of the most mature organizations have standardized all data management functions, whereas only 7% of the least mature organizations have standardized all data management functions.



Completed building single source of truth

Only 19% of organizations have completed building a single source of truth for all critical data domains. Over half of respondents haven't started yet or are still building and testing solutions. A single source of truth exists in the most mature organizations, but not in the least mature organizations.

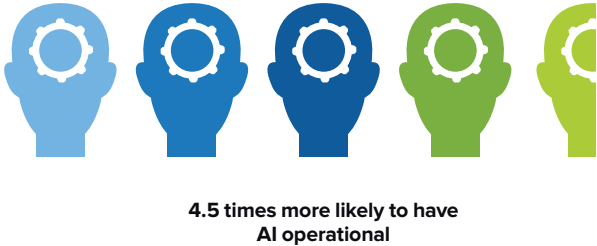


Over half haven't started yet or are still building and testing solutions

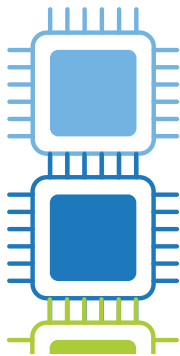
Sixty-eight percent of the most mature organizations have one central repository for all metadata, whereas only 5% of the least mature organizations have one repository.



The most mature organizations are 4.5 times more likely to have artificial intelligence (AI) operational across the organization compared with the least mature organizations.

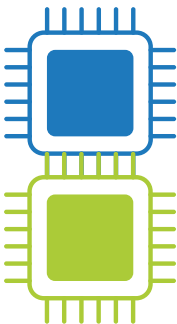


Ninety-five percent of the most mature organizations were mostly or completely satisfied with the value currently being realized from data management automation.



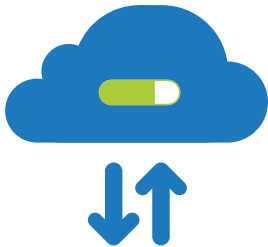
Mature 2.2 times more likely centralized data management

The most mature organizations are 2.2 times more likely to have centralized data management and two times more likely to have centralized analytics functions compared with the least mature organizations.



Mature 2 times more likely centralized analytics functions

Mature organizations always expect the availability of data intelligence when making a data-driven decision.



IDC Opinion

A digital-first world is emerging, and to be digital-first, organizations need to be data-driven and put data first. Putting data first will improve data-driven decision making and drive better business outcomes. But putting data first is difficult in modern data environments: The very things that are driving the need to put data first are the same things that are challenging the ability to move forward.

Modern data environments are characterized as having highly distributed, diverse, and dynamic data and data management technology. Data is distributed across hybrid and multicloud environments, stored and organized in many different formats and database management technologies, streaming at high frequencies, and constantly changing. The technology used to integrate and take control of data is also highly distributed, diverse, and dynamic. Organizations are dealing with a myriad of point solutions in organic environments that have grown out of the need to support legacy data in parallel with modern data in modern environments. Nearly 20% of an organization's data and analytics budget is being spent on integrating point solutions. A unified data platform allows for faster innovation with data-driven decisions while ensuring greater trust in the data that feeds analytic models and drives actions.

Organizations need to unify and simplify data management technologies to gain control over data in modern environments. This trend has become apparent in data architectures, such as data control planes, data fabrics, and meshes, and in the solutions that data integration and intelligence software vendors are bringing to market. To understand these trends and to get a sense of where organizations are when it comes to putting data first, Informatica commissioned a survey and study with IDC.

Becoming data-driven is a journey that varies across organizations, and this study was able to identify four different levels of maturity within the survey population. The study provides an understanding of what the most mature organizations are doing to put data first, contrasted against what the least mature organizations are doing across dimensions of data management functions, capture and curation of metadata, levels of automation, and approaches to decision making.

The most mature organizations identified in the study leverage a single platform for data management, share common metadata across all functions to deliver intelligence about data where and when it is needed, and embrace automation to deliver insights at scale. Organizations at the top of the maturity scale have experienced more than 2.5 times (+129%) the improvement in data metrics and a 107% improvement in business metrics compared with those at the bottom of the scale.

In This White Paper

This white paper describes highlights of an IDC global internet survey, commissioned by Informatica, on the topic of data management platforms. Survey respondents included senior professionals who are involved in management and investment decisions related to data and analytics management and governance. Demographics of survey respondents are available in the appendix. IDC was able to identify a maturity scale using positive correlations across dimensions of data management approaches, function implementation, degrees of automation, and how organizations approach the availability of data and intelligence about data to drive decision making. Four groups of respondents were created across the scale to enable the evaluation of characteristics that identify how organizations can improve data management maturity to deliver better business outcomes. Details on the method used to create the scale are available in the appendix.

Situation Overview

Regardless of where organizations are on their journey to becoming data-driven, there are some common characteristics exhibited by all, inclusive of organizational structure and priorities, current state of technology implementation in support of data management functions, use of metadata and automation, and how organizations are measuring success.

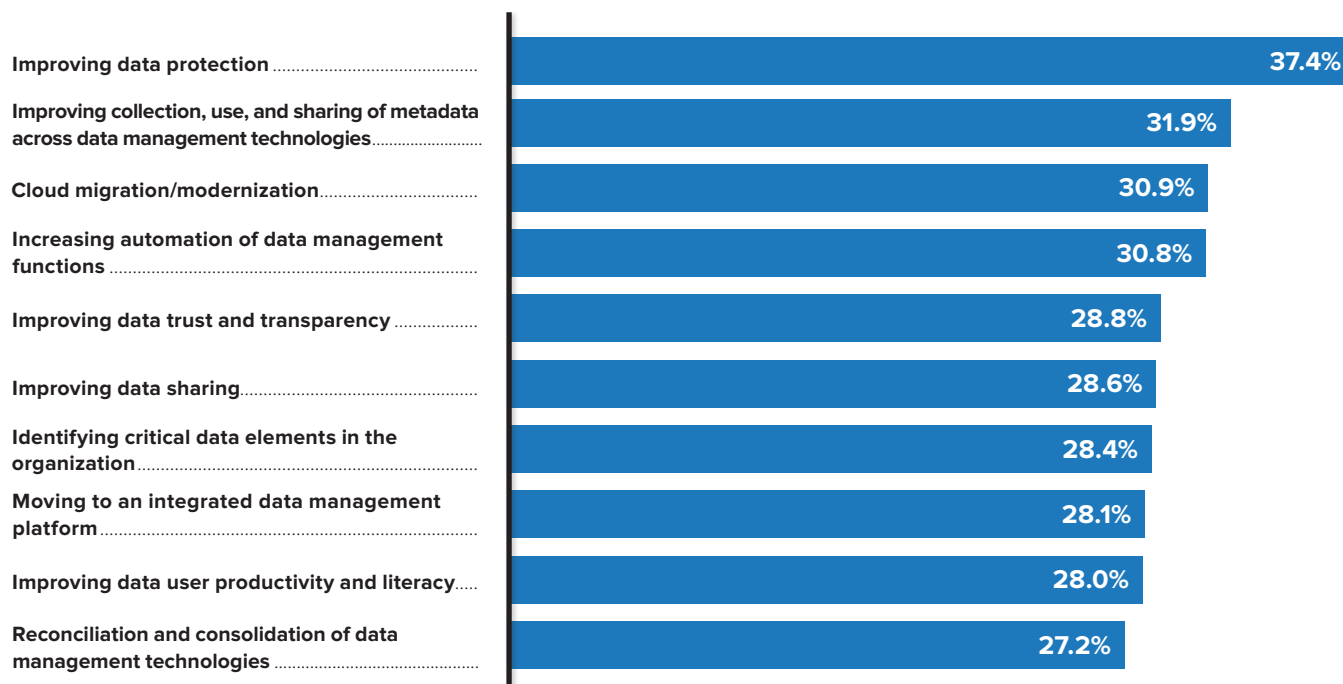
Organizational Structure and Priorities

Understanding the current state of data management includes looking at how organizations are structured. Data-driven organizations may separate data management accountability from analytics management accountability, and there are three primary models observed: centralized, decentralized, and embedded within the business. Organizations in the study showed similar ratios across the three models in both data and analytics management functions, with nearly half (49%) of the organizations having centralized structures, 40% having decentralized structures, and the remaining 11% having data management embedded in the business. Regardless of structure, the role of the chief data officer is becoming ever more important to ensure that data and analytics are being delivered consistently across the organization and that policies are in place to properly curate and protect data so it can be trusted.

Improving data protection is the highest area of focus for all organizations, followed closely by improving the collection, use, and sharing of metadata across data management technologies (see **Figure 1**, next page). In small-midsize (3,000 to 5,000 employees) and very large organizations (10,000+ employees), these priorities switch places, but only by a small margin. Tied for second and third place are cloud migration and increasing automation of data management functions. This insight is important because metadata provides the intelligence required to improve data protection and the use and level of trust in automation, and that technology in the modern environment must be cloud-native.

FIGURE 1**Areas in Data Management Needing Most Focus**

(% of respondents)

Q: Rank which areas need the most focus in addressing the improvements needed in managing data.

n = 819, Source: IDC's Global Data Management Platform Survey, 2022

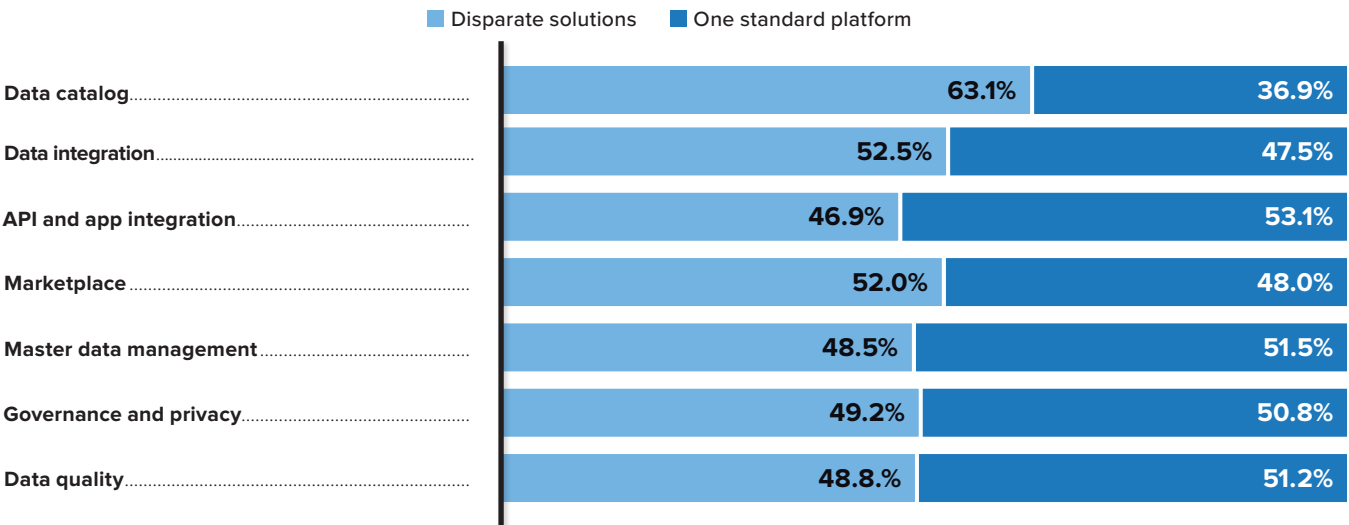
Current State of Data Management Technology

Data management technology as defined by this study is inclusive of data cataloging, data integration, application programming interface (API) and application integration, data quality, master data management, data governance and privacy, and data marketplaces. Data management technology enables organizations with data to improve the quality, trust, and protection of data to improve data analytics in support of data-driven decision making and outcomes. Data management technology makes data available to users, and yet only 58% of organizations believe that all or most of the data they have is available to be used by application developers, data stewards, data engineers, data analysts, data scientists, and business analysts. When data is available and accessible, companies can deliver insights at scale, which enables a sustainable competitive advantage and the ability to outperform the market. Only slightly more organizations (61%) believe the user experience in their current data management systems is meeting the needs of these same roles, and executives. This means there are still several gaps that need to be filled to make data more accessible and to improve the usability of data management technology.

There is approximately a 50:50 split between the number of organizations that are using manual, in-house-developed, or point solutions across all data management functions and organizations that have standardized on a data management or cloud platform solution. Only 16% of the total number of survey respondents have standardized all seven data management functions (see **Figure 2**).

FIGURE 2
Standardization of Data Functions
(% of respondents)

Q: Currently, how is the organization implementing the following data management functions?



Source: IDC's Global Data Management Platform Survey, 2022

When the individual functions are looked at in more detail, more organizations are using manual, in-house-developed, or point solutions for data cataloging, data integration, and data marketplaces than organizations that have standardized solutions. API and application integration, data quality, master data management, and governance and privacy functions have slightly more organizations using standardized solutions than those using point, in-house-developed, or manual solutions. Greater adoption of standardization can improve business planning, drive digital transformation initiatives, and add trust to data analytics and data decisions..

Metadata

A unifying component of data management functions is metadata: the data about data that answers the who, what, where, when, why, and how of data enablement. Metadata can be a unifying agent across disparate point solutions, and it usually is the foundational element of unified data platforms. Data intelligence is driven by metadata, and how an organization captures, curates, and uses metadata is a good indicator of

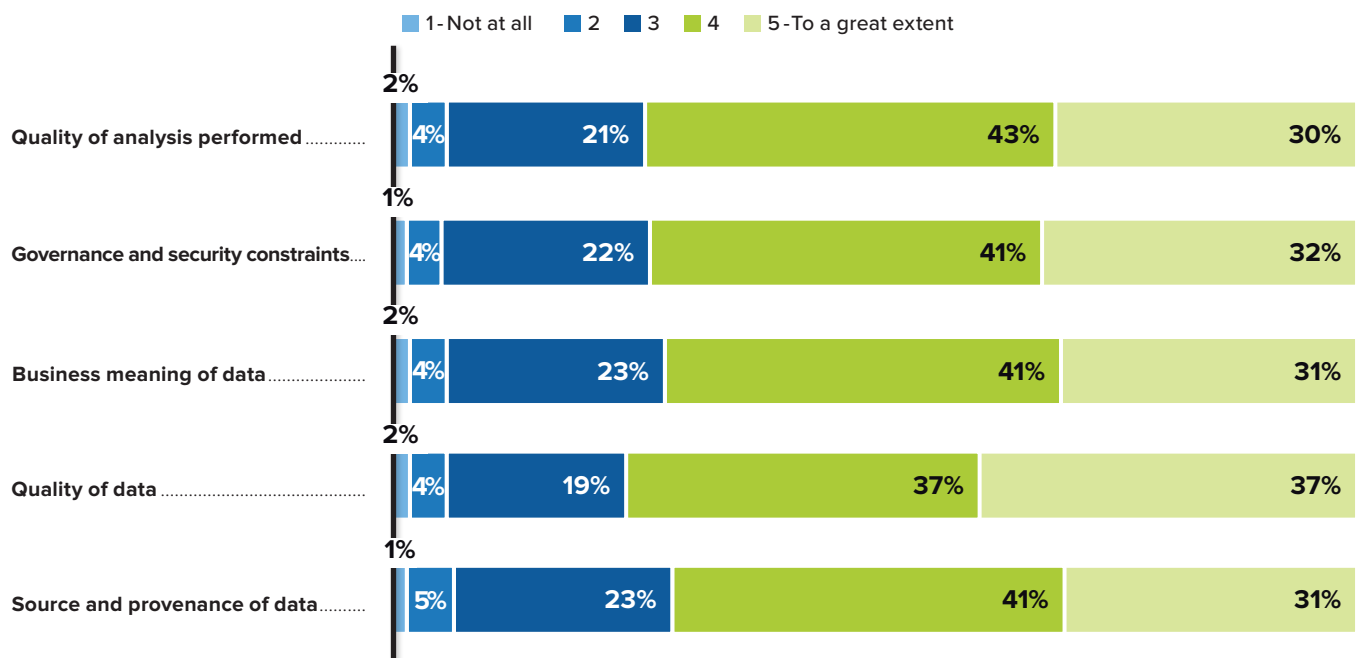
success toward becoming a data-driven organization, across the entire business. The organizations surveyed are almost equally divided between those that have an informal and siloed approach to metadata capture, curation, and sharing and those that have more formal processes. While formal versus informal methods of metadata capture and curation are split, only 31% of respondents have one centralized metadata repository. The whole of how organizations capture, curate, manage, and share metadata needs to be consistent in the drive toward putting data first.

Metadata is also what provides intelligence about data. Data intelligence answers the who, what, where, when, why, and how of data, and a unified approach builds trust in the data. Data intelligence is required to perform data governance processes, inform how data needs to be protected based on security and privacy classifications, identify high-value data, and put data into its proper context to ensure proper use. The data worker majority expects to have intelligence about data when making data-driven decisions, as illustrated in **Figure 3**.

FIGURE 3
Data Intelligence Expectations

(% of respondents)

Q: When you make data-driven decisions, to what extent do you expect and demand to know each of the following?



Source: IDC's Global Data Management Platform Survey, 2022

Automation

Automation in data management can be applied across all functions. Automation can also take many forms, including the use of rules and regular expressions to classify data, standardize data, weed out poor-quality data, and match and merge data. Automation can also leverage artificial intelligence and machine learning (ML) to identify anomalies, classify data, perform data discovery, make recommendations during development, and optimize runtime processes.

Automation can significantly improve decision making as well as reduce the amount of manual effort required in data management processes. It can reduce time to value by increasing discoverability and collaboration across business groups. As illustrated in **Figure 4**, data mastering, or master data management, is the one function that is the most manual compared with others. This could be a contributing factor to only 19% of the survey population having completed building a single source of truth for all critical data domains. Over half of the respondents haven't started yet or are still building and testing solutions.

FIGURE 4
Proportion of Manual Effort by Function
(% of respondents)

Q: What proportion of each data management function requires manual efforts?



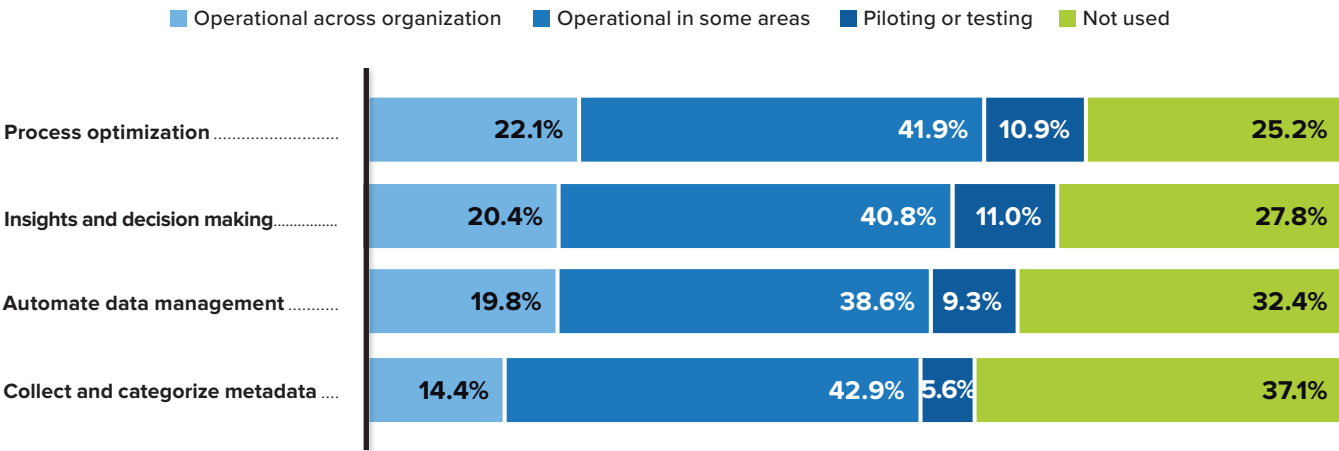
Source: IDC's *Global Data Management Platform Survey*, 2022

Figure 4 also indicates that data cataloging processes are the most automated. This is not surprising, given that data cataloging processes can be mechanical in the form of metadata crawlers and API connections across data and application processes. Data cataloging is also a good candidate for artificial intelligence and machine learning for data classification and discovery.

This finding leads into the question *How much is artificial intelligence being used across the organization?* Artificial intelligence is trained by machine learning. The study asked this question across four uses of AI: AI for insights and decision making, AI for data management automation, AI for process optimization, and AI for collecting and categorizing metadata. **Figure 5** illustrates that AI is being used the most for process optimization and automation, followed by finding insights and driving decisions, with automation of data and metadata management the least used. This is not surprising, as the early focus of AI was on finding insights for business decisions and less so on automating processes, including processes in the data organization. Use of AI in data management automation is growing in application within the data integration and intelligence software market.

FIGURE 5
Use of AI Automations
(% of respondents)

Q: AI can be applied to many processes in the organization. For each of these areas, how is AI currently being used?



Source: IDC's Global Data Management Platform Survey, 2022

Measurements of Success

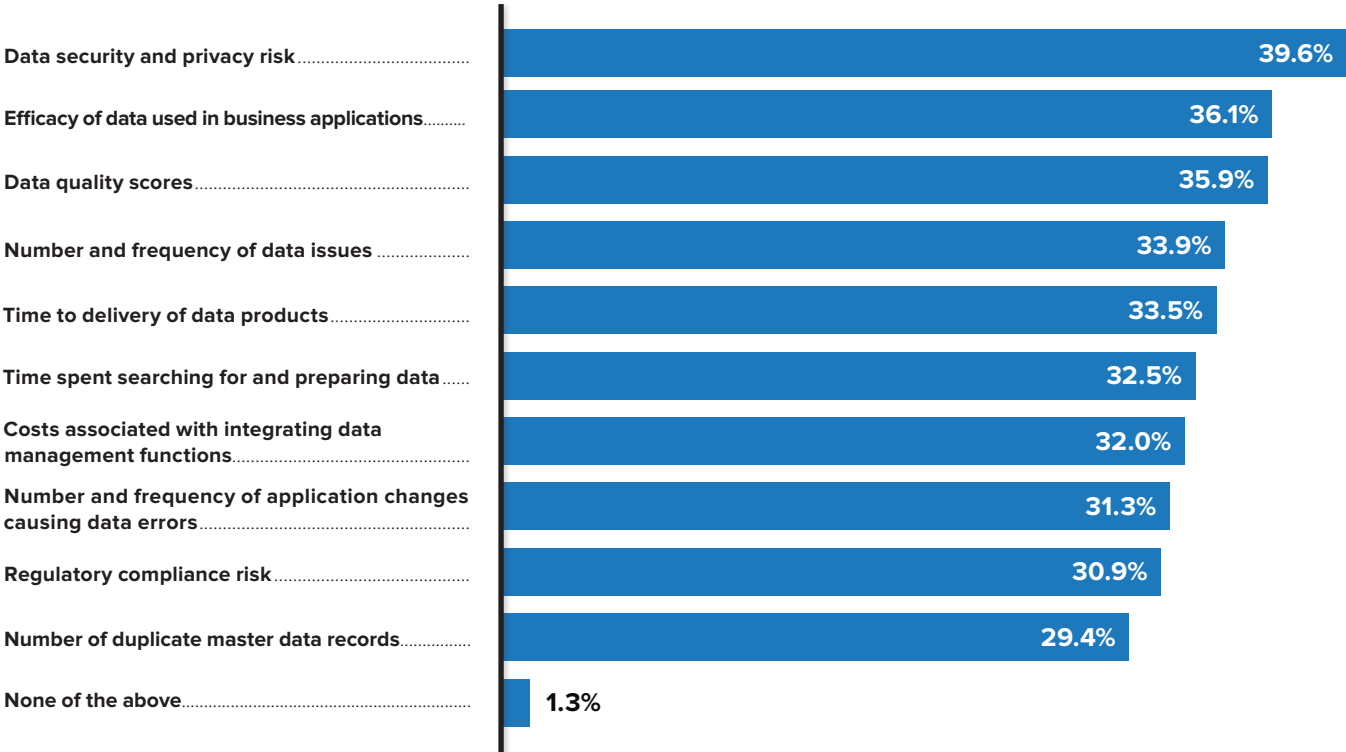
It has been said that improvements cannot be made if measurements aren't in place. Many different dimensions of data management processes and practices can be measured and monitored for improvements. **Figure 6** (next page) illustrates the percentage of organizations that are measuring each metric. The percentage of respondents measuring each metric is not high in the total population, which indicates there is room for improvement. But the fact that only 1.3% of the population is not measuring any of these metrics tells us that nearly all organizations are measuring something. Data security and privacy risk is being measured the most, followed by the efficacy of data used in business applications and data quality scores.

These results are consistent with other surveys IDC has done, where security and privacy are getting a lot of attention, as they should. But we also question at what cost, because tight controls can get in the way of innovation. Efficiency of data, AI/ML, and time spent searching all point to employee productivity and enterprise intelligence. Measuring data security, compliance, and data errors ensures the health of the business. Organizations need to find a balance to truly innovate with data. The fact that the number of duplicate master records is at the bottom of this list shows a strong correlation to other questions about challenges associated with implementing master data management solutions.

FIGURE 6
Data Metrics

(% of respondents)

Q: Which of the following data metrics are being measured to assess the success of your organization’s capabilities in data management?



Source: IDC’s Global Data Management Platform Survey, 2022

Resolving Complexity Through Maturity

Modern data environments are complex, as highly distributed, diverse, and dynamic data is being managed for the most part by standardized and non-standardized solutions within a mix of centralized and decentralized data management organizations. As not every organization is the same, IDC was able to build a maturity scale across the dimensions of:

- Organizational approach to implementing data management functions and management of metadata
- The degree of automation across data management function processes
- Data intelligence expectations, the current state of value being realized by automation, and the level to which automation could improve decision making

The scale follows a natural distribution, with all items positively correlated. The result is four individual groups of respondents at different levels of maturity across the scale:

- Reactive: The lowest 20th percentile of respondents on the scale
- Opportunistic: The 20th to 50th percentile of respondents on the scale
- Repeatable: The 50th to 80th percentile of respondents on the scale
- Optimized: The top 20th percentile (80–100) on the scale

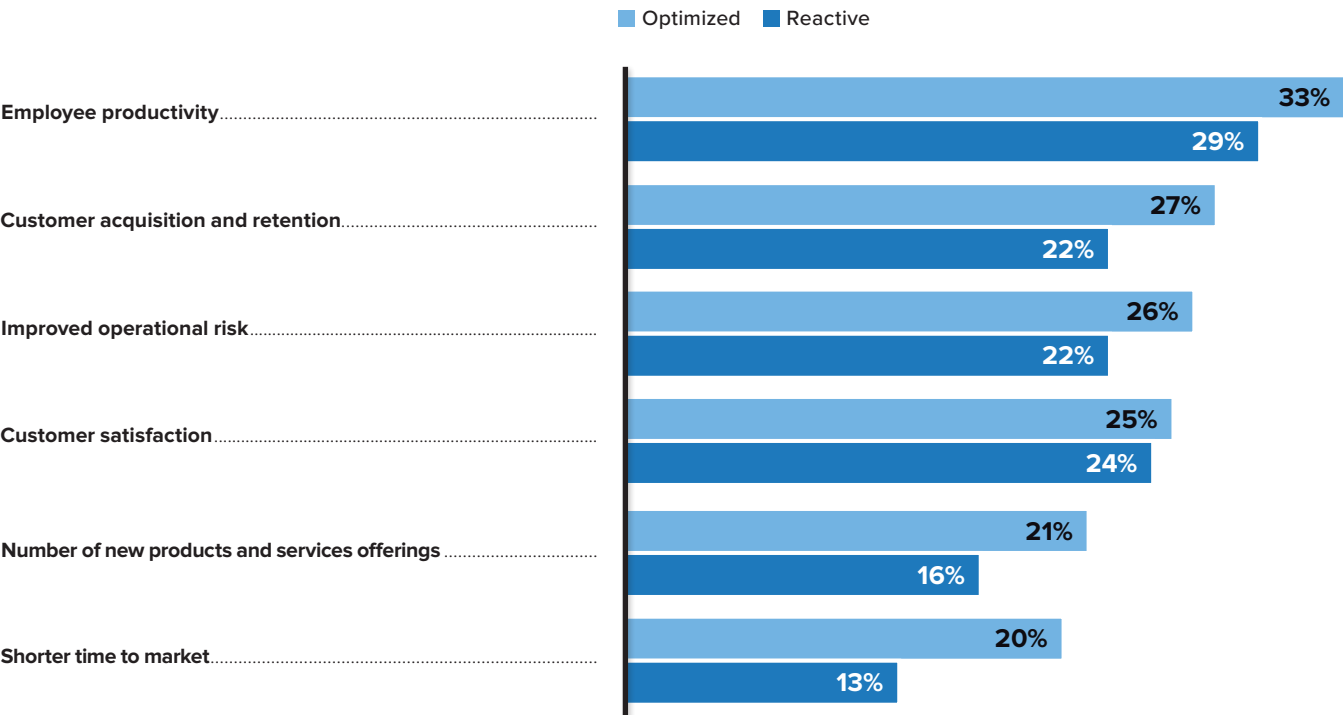
These groupings provide insight into what the least mature are doing compared with the most mature, while also providing insight into how organizations can progress across the levels by looking at what those in the middle are doing. This method helps answer the questions *Why is maturity important?*, *What are organizations at the top of the scale doing differently than those at the bottom?*, and *How are organizations achieving excellence?*

Why? Optimized Data Management Organizations Are Measuring Better Outcomes

Organizations at the top of the scale are seeing a combined improvement of 107% across all business metrics compared with those at the bottom of the scale. **Figure 7** illustrates the percentage of improvement that has been seen by organizations that have optimized data management compared with those that are reactive.

FIGURE 7
Business Outcomes
(% of improvement)

Q: As a result of investments in data management, how much has your organization improved its metrics for the following in the past two years?



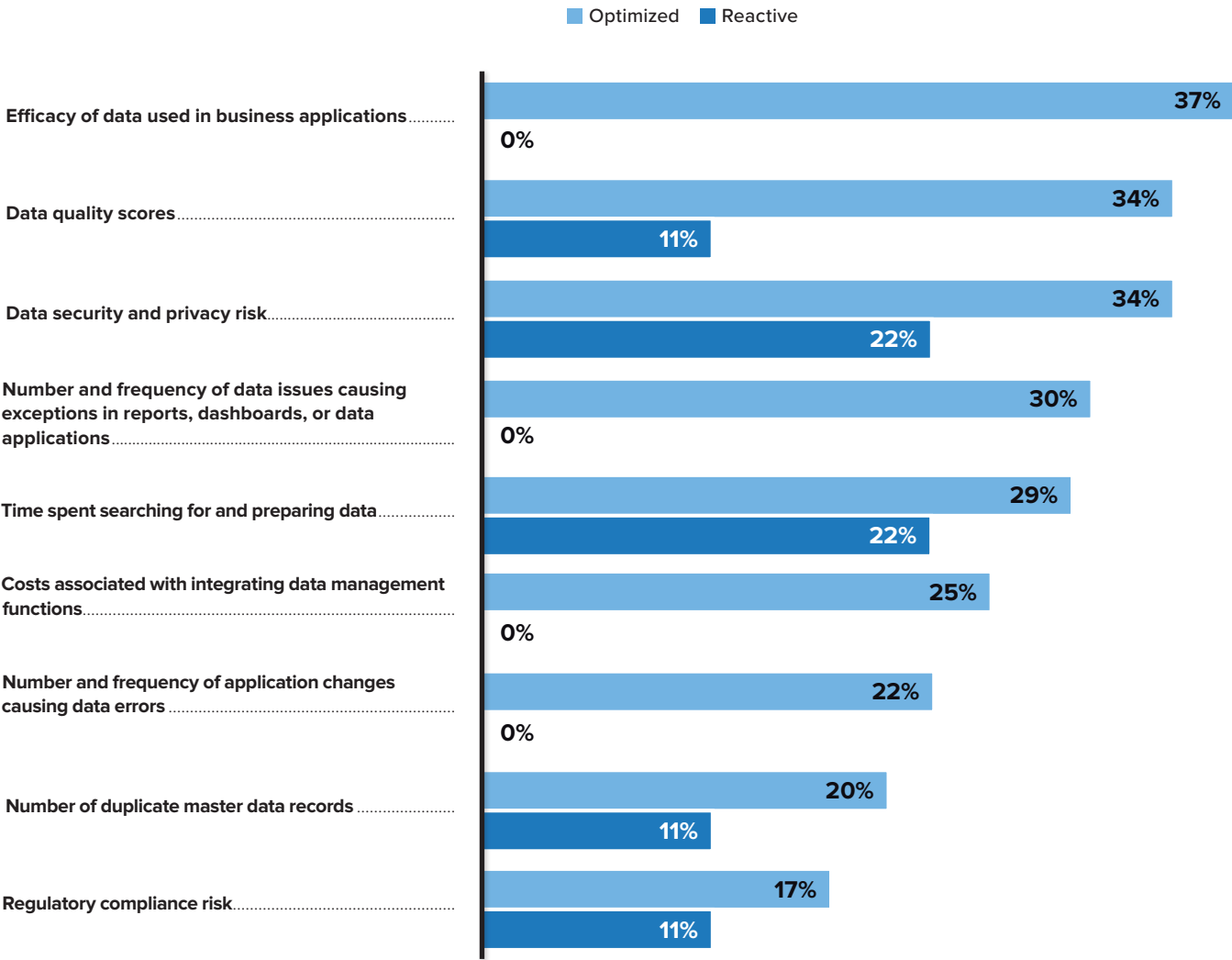
Source: IDC's Global Data Management Platform Survey, 2022

Organizations at the top of the scale are seeing even more improvements in data management metrics, at a factor of 2.2 times, or 117%. **Figure 8** (next page) illustrates the percentage of improvement of optimized and reactive organizations. It is important to point out that reactive organizations indicated that four of these metrics had zero improvement: efficacy of data, data exceptions, data management function integration costs, and frequency of application changes causing data errors. Optimized organizations are also measuring twice as many metrics as reactive organizations, but even the metrics that are reported as having zero improvement are being measured by some reactive organizations.

FIGURE 8
Data Management Outcomes

(% of improvement)

Q: Over the past two years, how have your organization’s capabilities in data management impacted each of these metrics?



Source: IDC's Global Data Management Platform Survey, 2022

What? Optimized Data Management Organizations Have Standardized Data Management Functions and Embraced Automation

Optimized organizations are on average **1.7 times** more likely to use one standardized platform for data management functions compared with reactive organizations.

Table 1 provides the multiple by function.

TABLE 1
Use of One Platform: Optimized over Reactive

Q: Currently, how is the organization implementing the following data management functions?

Function	Number of Optimized Over Reactive Organizations Multiple
Data cataloging	2.2
Data integration	1.3
API and application integration	1.6
Data quality	1.6
Master data management	1.7
Governance and privacy	1.6
Data marketplace	1.6

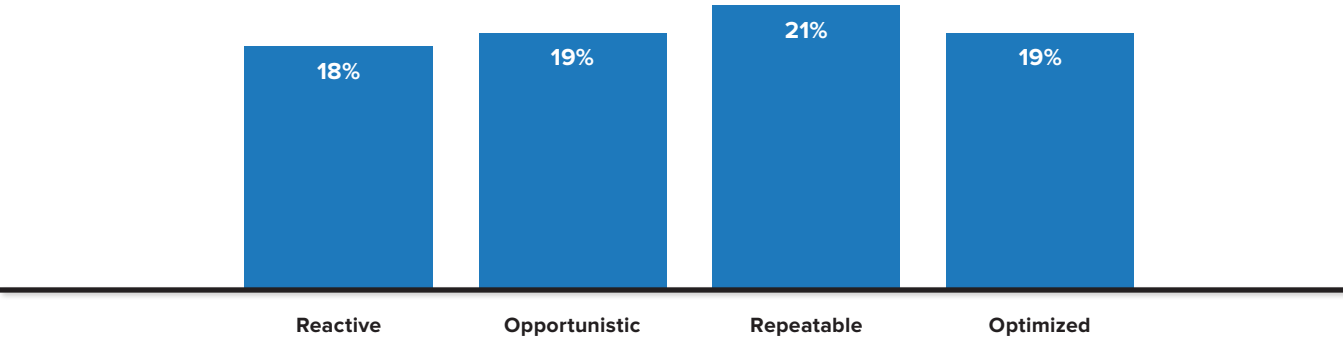
Source: IDC's Global Data Management Platform Survey, 2022

Forty percent of optimized organizations have standardized all functions, whereas only 7% of reactive organizations have standardized all functions. Therefore, optimized organizations are **nearly six times** (5.7%) more likely to standardize all functions. The survey instrument did not provide an understanding of whether all seven functions were standardized on the same platform.

Figure 9 illustrates the proportion of an organization's spend on integrating different point solutions into the data management platform. The largest percentage of spend is happening in the third category on the scale, the organizations classified as repeatable. The percentage of spend is lowest in the reactive organizations, in part because they do not have many of the functions and likely have not spent time integrating functions. Spend goes up as you move across the scale, but then drops back down for the optimized organizations. Here again, we can see that even if the optimized organizations have standardized functions, not all functions may be on the same platform. Spend for the organizations in the repeatable category is the highest, which is likely why moving to an integrated data management platform is also a big focus for organizations in this category. Repeatable, autonomous operations — enabled by data governance and standardization — allow data-driven organizations to scale up more business use cases and realize value from digital transformation.

FIGURE 9
Budget Spend on Integration by Category
(% of budget spend on integration)

Q: Of the total spend on data management and analytics, what proportion of your organization's spending is on integrating different point solutions into the overall data management platform?



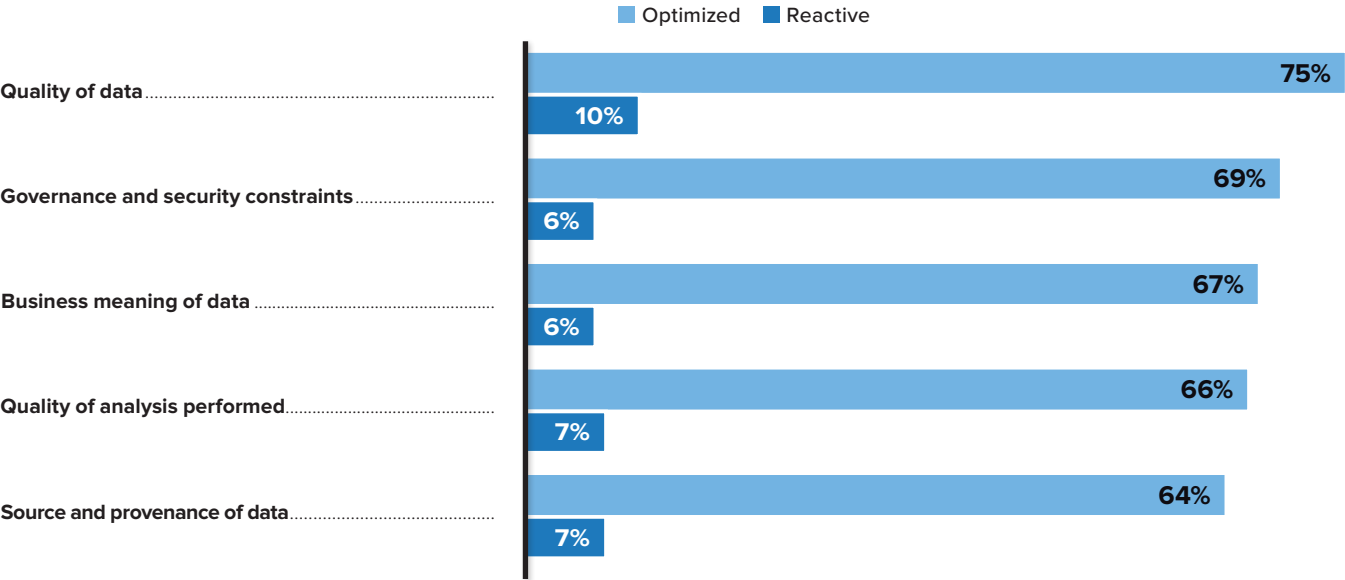
Source: IDC's Global Data Management Platform Survey, 2022

Optimized organizations also manage metadata much differently than reactive organizations. Sixty-eight percent of optimized organizations have one central repository for all metadata, whereas only 5% of reactive organizations have one repository. Most of the reactive organizations have multiple silos of repositories or take a federated approach. When it comes to the expectations of having intelligence about data available to decision makers, **Figure 10** (next page) illustrates the significant difference between optimized and reactive organizations in terms of the expectations around the availability of each dimension of intelligence when making data-driven decisions.

FIGURE 10
Data Intelligence Expectations

(% of respondents who expect intelligence to a great extent)

Q: When you make data-driven decisions, to what extent do you expect and demand to know each of the following?



Source: IDC's Global Data Management Platform Survey, 2022

Figure 10 illustrates that optimized organizations are leveraging and activating data intelligence to inform better decision making, contributing to the improved business and data management outcomes reported.

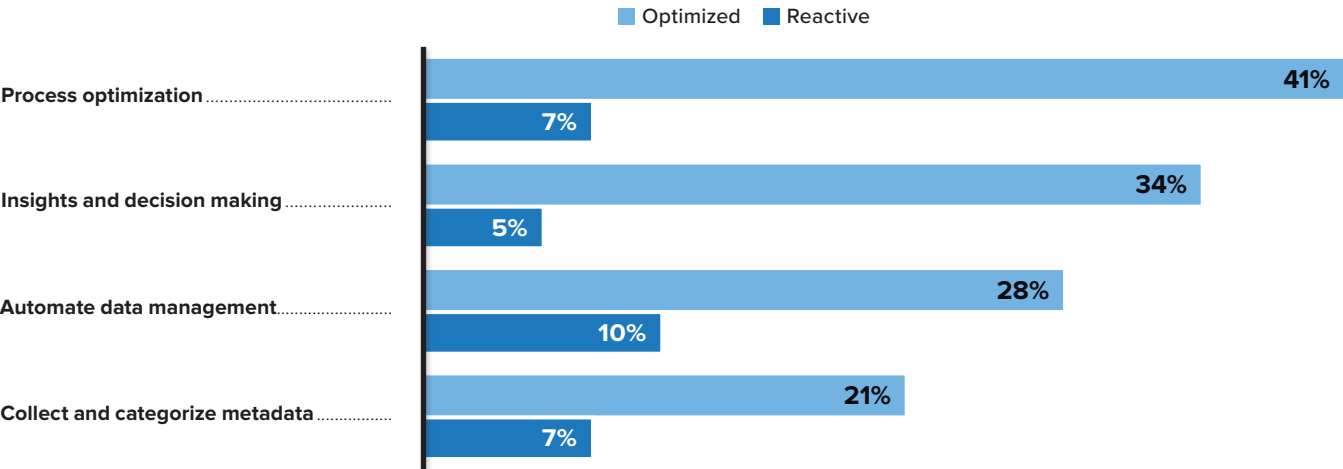
Optimized organizations are also much more likely to embrace automation.

Figure 11 (next page) illustrates the percentage of optimized organizations that have AI operational across the organization compared with reactive organizations that have the same level of AI deployed. Optimized organizations are on average 4.5 times more likely to have AI operational across the organization compared with reactive organizations. Removing time-consuming manual tasks associated with data intelligence while deploying AI allows a company to immediately make data-driven decisions for both internal processes and external customer experiences, without delay.

FIGURE 11
Deployment of AI Automation

(% of respondents with AI operational across the organization)

Q: For each of these areas, how is AI currently being used?



Source: IDC's Global Data Management Platform Survey, 2022

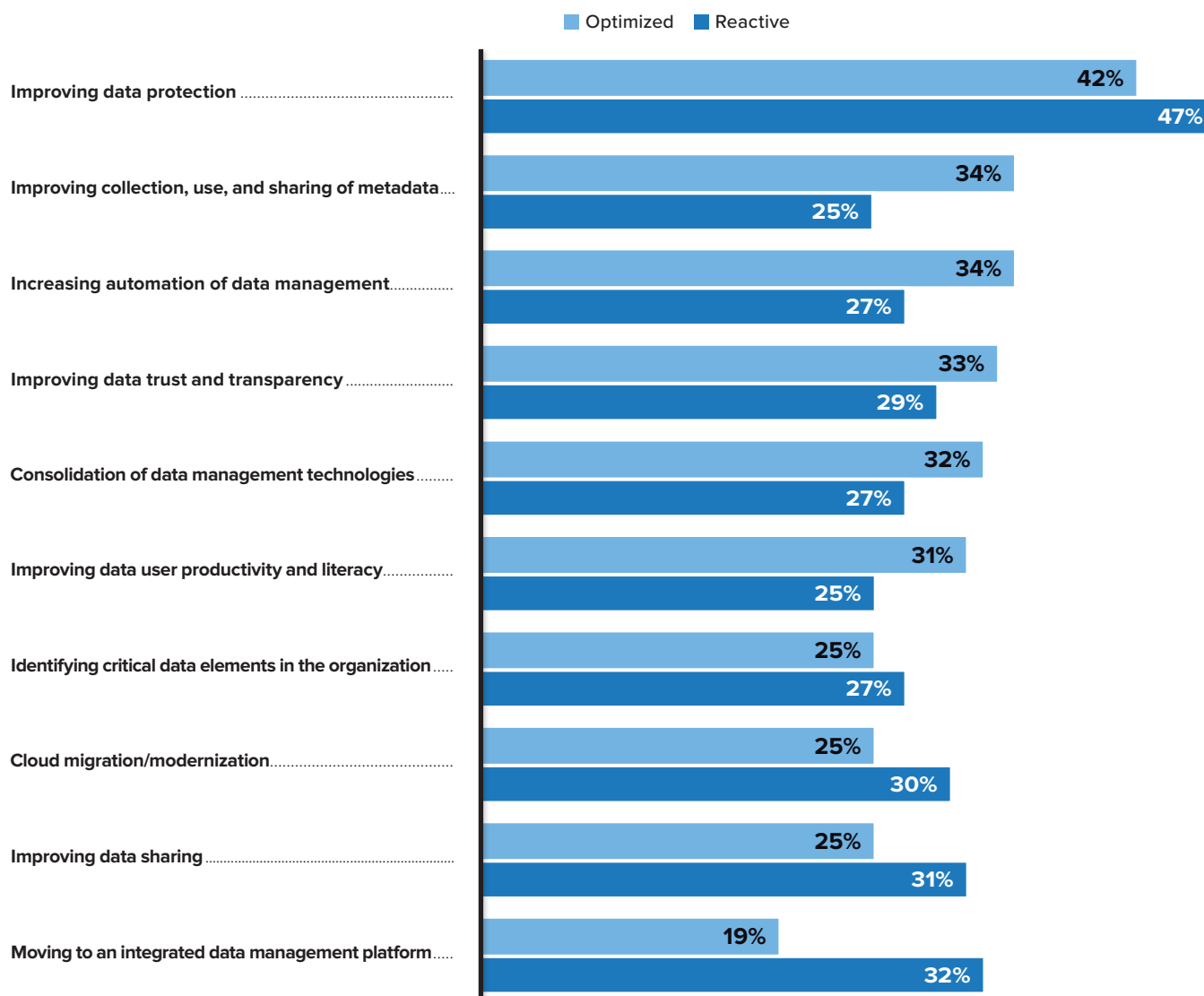
How? Optimized Organizations Are Setting Many Objectives and Efficiently Manage and Share Data, with Centralized Organizational Structures That Can Drive Business Value

How organizations approach the drive to becoming data-driven is reflected in the focus each has on specific data management objectives, how much data is available to different roles within the organization, and how the data needs of different personas are being met. All of these and the organizational structure of the data management team can also provide a glimpse into an organization's data culture.

Figure 12 (next page) illustrates the percentage of optimized organizations that have a specific focus on each of the stated areas in data management compared with reactive organizations. Improving data protection is always at the top of the list in areas of focus in data management, but there are fewer optimized organizations that have this objective compared with reactive organizations. This illustrates that perhaps data protection is a capability that optimized organizations have achieved. Next to data protection, optimized organizations have a focus on metadata and increasing automation. Next to data protection, reactive organizations have a focus on moving to an integrated data management platform, improving data sharing, and cloud migration. We can also see that fewer optimized organizations are focused on moving to one integrated data management platform because we have already established that some already have one.

FIGURE 12**Data Management Focus Areas**

(% of respondents)

Q: Rank which areas need the most focus in addressing the improvements needed in managing data.

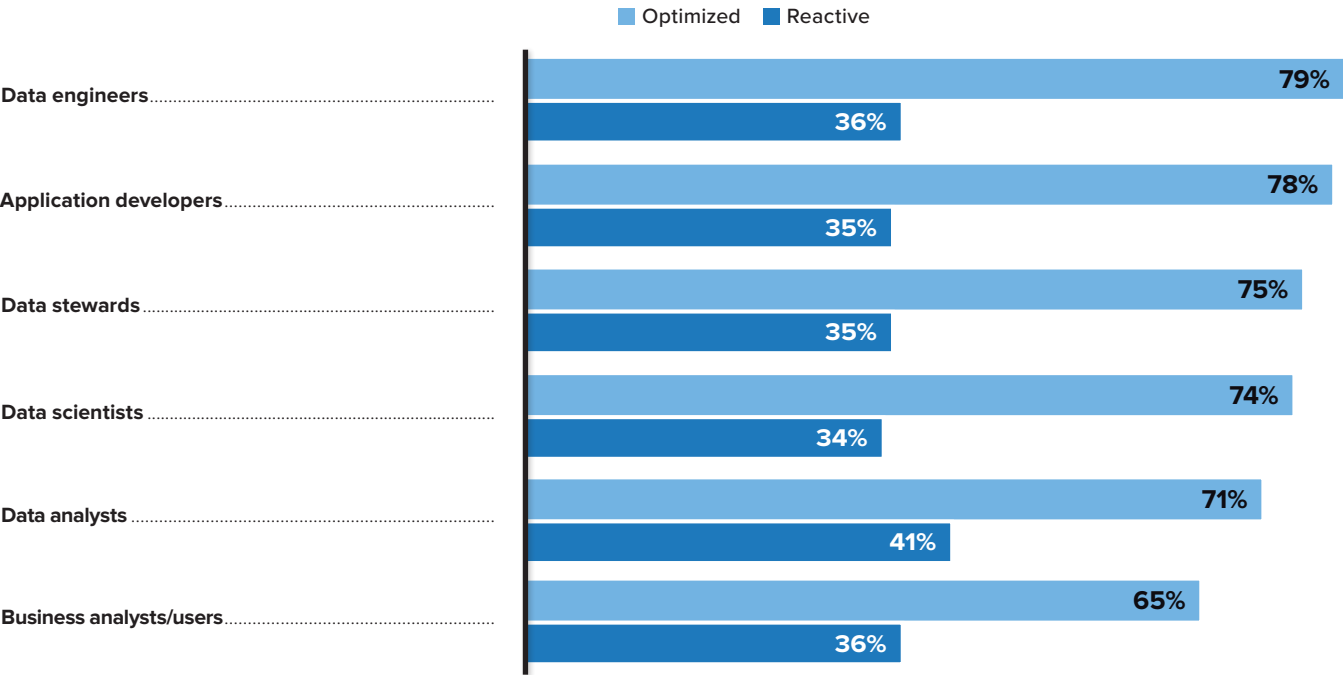
Source: IDC's Global Data Management Platform Survey, 2022

Optimized organizations are on average twice as likely to have all or most data available to personas within the organization. **Figure 13** (next page) illustrates the dramatic difference at both ends of the scale, for each persona. Data engineers are at the top of the list of roles that optimized organizations make data available to, whereas data analysts are at the top of the list for reactive organizations. This indicates that reactive organizations may not yet be at the point where data engineering has become a discipline.

FIGURE 13
Data Availability by Role

(% of respondents with all or most data available)

Q: What degree of data is readily and easily available for each of these roles in your organization?



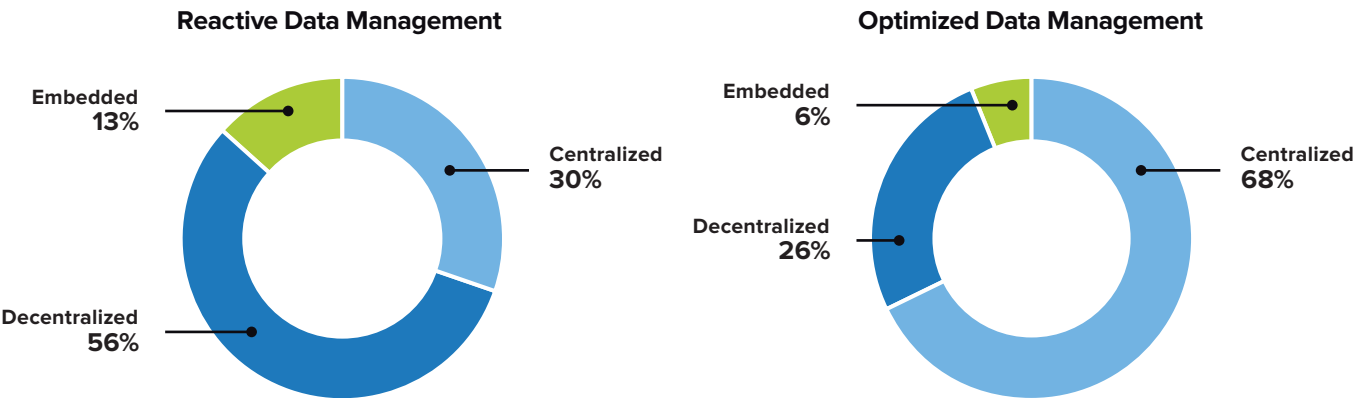
Source: IDC's Global Data Management Platform Survey, 2022

Optimized organizations are also **2.3 times, or 130%**, more likely to be meeting user experience requirements with the data management systems currently deployed across the same set of personas compared with reactive organizations. Requirements are inclusive of recommendations, guided workflows, and automation, each of which requires some level of intelligence in the system. When asked how satisfied the organization is with the value currently being realized from data management automation, **95%** of optimized organizations were mostly or completely satisfied. Only 22% of reactive organizations felt the same way. Although more of the optimized organizations are satisfied with the value of intelligent automation, nearly the same percentage (93%) believed that there is room for improvement in the use of automation to improve business decision making. Only 29% of reactive organizations see the need to improve the use of automation. The net of this finding: Reactive organizations do not know what they do not know.

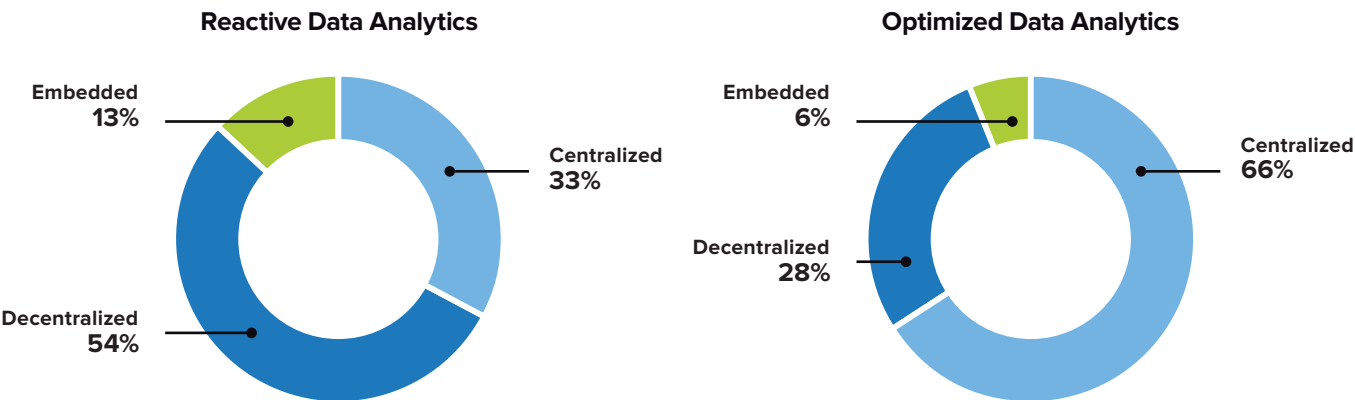
The final characteristic that provides insight into how optimized organizations are successful in data management looks at the data and analytics organizational structures. **Figure 14** illustrates the difference in centralized versus distributed structures of the data management and data analytics functions. Optimized organizations are 2.2 times more likely to have centralized data management and two times more likely to have centralized analytics functions compared with reactive organizations. Twice as many optimized organizations have centralized structures, and more than half as many have decentralized and business-embedded resources.

FIGURE 14
Data and Analytics Organizational Structures
(% of respondents)

Q: Which statement best describes how the data management function is organized?



Q: Which statement best describes how the data analytics function is organized?



Source: IDC's Global Data Management Platform Survey, 2022

Future Outlook

As the modern data environment continues to grow in complexity, organizations will need to make further improvements and investments in technology selection and implementation, metadata management and activation, use of automation, and meeting the needs of a diverse set of data worker personas. Establishing technology platforms to boost enterprise intelligence will provide a sustainable competitive advantage and the ability to outperform the market.

Technology

- As this study indicates, organizations that are getting better business and data management outcomes have standardized data management functions. Although it is not clear from the study that all functions have been standardized on the same platform in every optimized organization, there are indications that some have standardized on a single platform. One single platform reduces the time and money required to integrate disparate point solutions, regardless of those solutions being commercial, open source, or bespoke.
- Metadata is a foundational component in the highly distributed, diverse, and dynamic modern data environment. Organizations have attempted to resolve the challenges of modern data environments by centralizing all of the data into data lakes and data warehouses. That can be part of the solution, but a better solution, represented in the results of this survey, is to centralize the metadata — the intelligence about the data. Once the metadata is centralized, it can be activated in the implementation of intelligent data management automation.

Automation

- Automation is happening today, but more can be done. Even respondents in this study who are part of the most optimized organizations see room for improvement. Disparate data management technologies and siloed metadata are holding back the opportunity for more automation in data management. The more that each data management function knows about what is happening in other data management functions, the more each function can learn from and leverage the intelligence of adjacent functions.

Organization

- Clearly, centralized data and analytics organizational structures align with better business and data management outcomes. As the world has changed, it is important to understand that *centralized* in this sense does not mean physically but logically. Similarly, there is a lot of value in combining business knowledge with data and analytics. Here is where, at a minimum, centralized command and control of data and analytics functions can be in place, regardless of where a data worker resides within the organization. Central data and analytics management policies, processes, and authority will be important in the digital-first world, helping organizations put data first.

Challenges/Opportunities/ Recommendations

Every organization represented in this study has challenges; the difference lies in whether or not the challenges are known and how these challenges are perceived by the organization. There are varying degrees of focus across the spectrum of improvements in data management to meet the challenges of modern data environments (refer back to **Figure 12**).

Challenges can become opportunities, and the data illustrates that optimized organizations are not focused on the same things that reactive organizations are if we put aside the need to protect data (which is usually at the top of the list). Furthermore:

- **Consider your company's data culture.** Strong data cultures deliver better business outcomes. Data-driven organizations are better at capturing knowledge, synthesizing information, and delivering insights at scale. Optimized organizations are more focused on strategic initiatives — such as improving the use of metadata, improving data trust and transparency, improving data literacy and user productivity, reconciling and consolidating data management technologies, and increasing data management automation — compared with reactive organizations, which are more focused on tactical initiatives.
- **Make the business case for optimization and automation.** Organizations at the top of the maturity scale see a combined improvement of 107% across all business metrics compared with those at the bottom of the scale. Instead of making a rushed decision about technology investment, take a long-term view of what will aid in future agility and resiliency. The right technology investments are required to remain competitive and have real bottom-line impact.
- **Be deliberate about moving up the maturity scale.** An interesting insight that this study provides is that as organizations move across the maturity scale, focus areas shift. Set proper expectations based on budgets and organizational resources. Organizations in the opportunistic category (20th to 50th percentile) share similar focus areas with the reactive organizations, but organizations in the repeatable category (50th to 80th percentile) have some very different areas of focus compared with optimized organizations. Those in the repeatable category are most focused on improving metadata collection, migrating to cloud, increasing automation, and moving to an integrated data platform as

compared with the other levels. This validates that the path to becoming optimized in data management is a journey, where focus and priorities will shift. Once you are optimized, you can enable your organization with data, which leads to faster insights, cost reductions, and higher levels of customer acquisition.

As you move up the maturity scale, consider this “best practices” checklist to assess and track your progress:

- **Standardize data functions.** Greater adoption of capabilities standardized on one platform can improve business planning, drive digital transformation initiatives, and add trust to data analytics and data decisions.
- **Enable data automation.** The scale of data distribution, diversity, and dynamics in modern data environments renders manual processes ineffective and obsolete.
- **Measure data management impact.** The efficiency of data, AI/ML, and time spent searching all point to employee productivity and enterprise intelligence. Measuring data security, compliance, and data errors ensures the health of the business.
- **Improve data management maturity.** Improvements in maturity levels are positively correlated with business outcomes: Improve your data management maturity and improve your top and bottom lines.
- **Accelerate time to value with a unified data platform.** Organizations at the top of the maturity scale are on average 1.7 times more likely to use one standardized platform for data management functions, realizing operational, resource, and processing efficiency benefits over those at the bottom of the scale that are leveraging multiple point solutions.
- **Optimize spend, time, and resources.** Removing time-consuming manual tasks associated with data integration and intelligence by deploying AI allows a company to immediately make data-driven decisions for both internal processes and external customer experiences, without delay.
- **Make your data accessible.** Organizations at the top of the maturity scale are, on average, twice as likely to have all or most data available to data-oriented personas within the organization compared with organizations at the bottom of the scale.

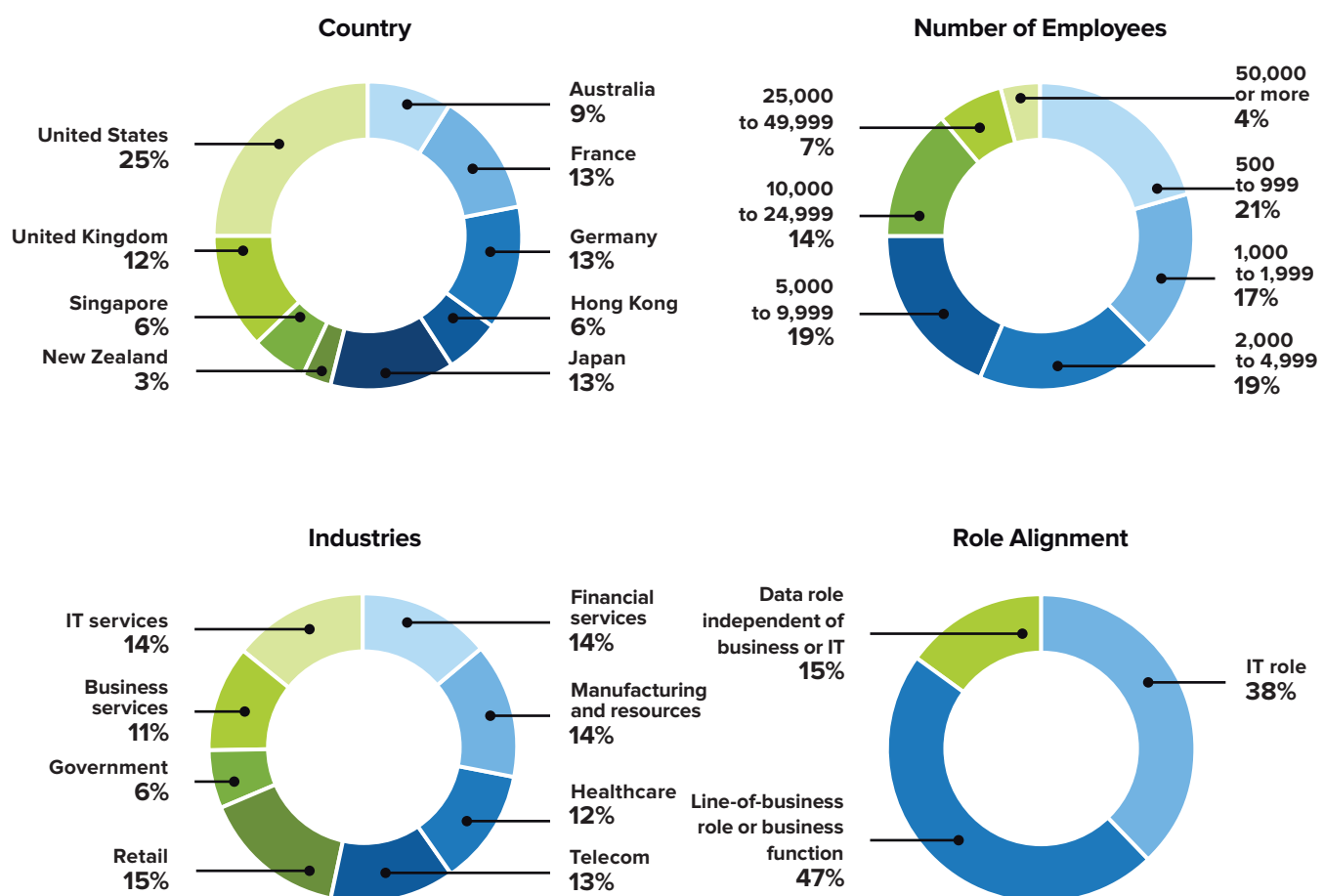
Conclusion

Becoming data-driven in a digital-first world requires organizations to put data first. Putting data first requires a focus on technology reconciliation and consolidation to work toward one platform, centralized and shared metadata, intelligent automation, and an effective data and analytics management organization to provide centralized command and control of data activities to improve trust and business outcomes.

Appendix

Survey Demographics

FIGURE 15
Study Demographics
(% of respondents)



Source: IDC's Global Data Management Platform Survey, 2022

Building the Maturity Scale

In building the data management maturity scale, the goal is to score respondents from 0 to 100, with inputs derived from attributes that measure their maturity or approach to data management.

The attributes that measure data management are split into three subscales in the survey instrument:

- **Section B:** Data management
- **Section K:** Data management approaches and decision making
- **Automation:** Use of automation across data management functions

The IDC approach is to build subscale scores for each respondent, measuring each of these areas on a 0–100 basis and combining them to form the final score.

In building each subscale, we look for survey results that measure the construct: a single variable in a section that best represents maturity within that area. Choosing single variables helps IDC in designing an easy-to-administer typing tool after final approval of the scale.

We evaluate the candidate variables using two criteria:

- **Face validity:** Does it make sense to include the variable in the subscale?
- **Construct validity:** Is the variable included in the subscale measure a good representation for other variables in the area, and are these variables significantly associated with each other? Does the variable impact business outcomes?

In the overall scale, we are looking for subscales that, when combined, tell where a respondent lands on the data intelligence scale. The approach to evaluate it is similar. We evaluate the candidate variables with the following criteria:

- **Face validity:** Does it make sense to include the subscale variable in the overall scale?
- **Construct validity:** Are the subscales significantly associated with each other? Does the score on the scale impact business value and other variables not included in the model?

Attributes to measure subscales across the six sections in the survey used at least one question from each section.

Each subscale was recoded to be on a 0–100 scale. The overall scale is calculated to be the average score across all the subscales (equal weighting). The result is a standard distribution of respondents to the survey. This distribution was then divided into four categories based on the percentile, to separate respondents across the scale into unique categories to observe characteristics and best practices across the scale.

About the Analyst



Stewart Bond

Research Director, Data Integration and Data Intelligence Software, IDC

Stewart's core research coverage includes watching emerging trends that are shaping and changing data movement, ingestion, transformation, mastering, cleansing, and consumption in the era of digital transformation. Having worked in the IT industry for over 25 years, from early experience in database and application development through solution design and deployment to strategic architectural consulting, Stewart has worked through some significant changes in the IT industry. His depth of field experience coupled with market insight gives him a unique perspective, valued by his customers and peers.

[More about Stewart Bond](#)



This publication was produced by IDC Custom Solutions. As a premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets, IDC's Custom Solutions group helps clients plan, market, sell, and succeed in the global marketplace. We create actionable market intelligence and influential content marketing programs that yield measurable results.



@idc



@idc

idc.com

© 2022 IDC Research, Inc. IDC materials are licensed for external use, and in no way does the use or publication of IDC research indicate IDC's endorsement of the sponsor's or licensee's products or strategies.

[Privacy Policy](#) | [CCPA](#)