Co-sponsored by: zoominfo

BEST PRACTICES REPORT

Q1 2024

Reducing Time to Insight and Maximizing the Benefits of Real-Time Data

By David Stodder

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

Reducing Time to Insight and Maximizing the Benefits of Real-Time Data

By David Stodder

Table of Contents

Executive Summary 5

Streamlining the Path to Data Insights 6

- Overall Satisfaction and Business Benefits Sought 8
- Modernization Priorities for Faster Data Insights 10

Data Currency and Real-Time Data 11

- Data Currency Types Available to Users . . . 13
- Issues Hindering Timely Data Insights 14

Data Management Strategies for Reducing Latency 16

- Addressing Data Integration and Preparation Challenges 19
- Expanding Monitoring to Holistic Data Observability 20

Importance of Data Intelligence and Data Catalogs 21

- Data Intelligence Modernization Priorities . 23

AI-Driven Automation and Augmentation 25

Priorities for Improving User Experiences 27

Recommendations 29

Research Co-sponsor: ZoomInfo 32

© 2024 by TDWI, a division of 1105 Media, Inc. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. Email requests or feedback to info@tdwi.org.

Product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies. Inclusion of a vendor, product, or service in TDWI research does not constitute an endorsement by TDWI or its management. Sponsorship of a publication should not be construed as an endorsement of the sponsor organization or validation of its claims.

This report is based on independent research and represents TDWI's findings; reader experience may differ. The information contained in this report was obtained from sources believed to be reliable at the time of publication. Features and specifications can and do change frequently; readers are encouraged to visit vendor websites for updated information. TDWI shall not be liable for any omissions or errors in the information in this report.

About the Author

DAVID STODDER is senior director of TDWI Research for business intelligence. He focuses on providing research-based insights and best practices for organizations implementing BI, analytics, data discovery, data visualization, performance management, and related technologies and methods and has been a thought leader in the field for over two decades. Previously, he headed up his own independent firm and served as vice president and research director with Ventana Research. He was the founding chief editor of *Intelligent Enterprise* where he also served as editorial director for nine years. You can reach him by email (dstodder@tdwi.org), on Twitter, and on LinkedIn.

About TDWI Research

TDWI Research provides industry-leading research and advice for data and analytics professionals worldwide. TDWI Research focuses on modern data management, analytics, and data science approaches and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of business and technical challenges surrounding the deployment and use of data and analytics. TDWI Research offers in-depth research reports, commentary, assessments, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.

About the TDWI Best Practices Reports Series

This series is designed to educate technical and business professionals about new business intelligence technologies, concepts, or approaches that address a significant problem or issue. Research for the reports is conducted via interviews with industry experts and leading-edge user companies and is supplemented by surveys of business intelligence professionals.

To support the program, TDWI seeks vendors that collectively wish to evangelize a new approach to solving business intelligence problems or an emerging technology discipline. By banding together, sponsors can validate a new market niche and educate organizations about alternative solutions to critical business intelligence issues. To suggest a topic that meets these requirements, please contact TDWI Senior Research Directors David Stodder (dstodder@tdwi.org), James Kobielus (jkobielus@tdwi.org), or Fern Halper (fhalper@tdwi.org).

Acknowledgments

TDWI would like to thank many people who contributed to this report. First, we appreciate the many users who responded to our survey, especially those who agreed to our requests for phone interviews. Second, our report sponsors, who diligently reviewed outlines, survey questions, and report drafts. Finally, we would like to recognize TDWI's production team: James Powell, Lindsay Stares, Pete Considine, Rod Gosser, and John Bardell.

Sponsors

Denodo, SAP, Snowflake, and ZoomInfo sponsored the research and writing of this report.

Report purpose. Organizations across industries are pursuing faster speed to insight to compete more effectively, to run operations more efficiently, to innovate, and to drive analytics and AI/ML for smarter customer relationships. Data is getting faster; real-time data sources are exploding. Bottlenecks persist in traditional systems and architectures. Organizations can take advantage of modern data platforms, data integration, and data intelligence to streamline processes and accelerate insights. This TDWI Best Practices Report examines priorities and offers best practices recommendations.

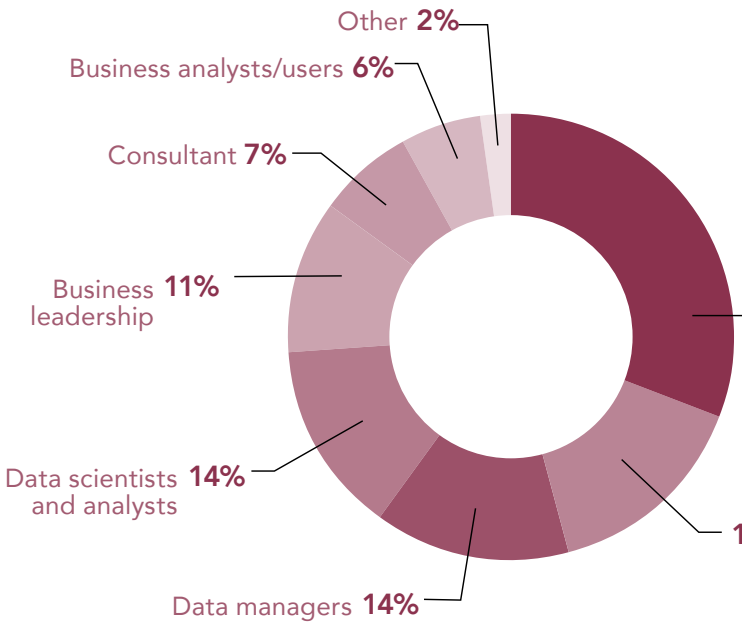
Survey methodology. In December 2023 and January 2024, TDWI sent invitations via email to business and IT professionals in our database, asking them to participate in an internet-based survey. The invitation was also posted online and in publications from TDWI and other firms. The survey collected responses from 257 respondents, with 181 completing every question. For this research, all responses are valuable and are included in this report's sample. This explains why the number of respondents varies per question.

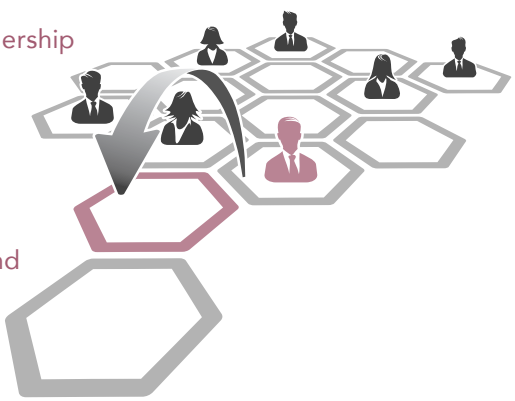
Research methods. In addition to the survey, TDWI conducted interviews with business and IT executives and managers, application developers, and data management and analytics experts. TDWI also received briefings from vendors that offer products and services related to the topics addressed in this report.

Survey demographics. Nearly one-third of respondents (31%) were in IT leadership positions such as CIO, CDO, and IT manager. These were followed by architects and developers (15%), data managers (14%), and data scientists and analysts (14%). Business leadership, analysts, and users made up a combined 17% of respondents, with consultants and other positions making up the remainder.

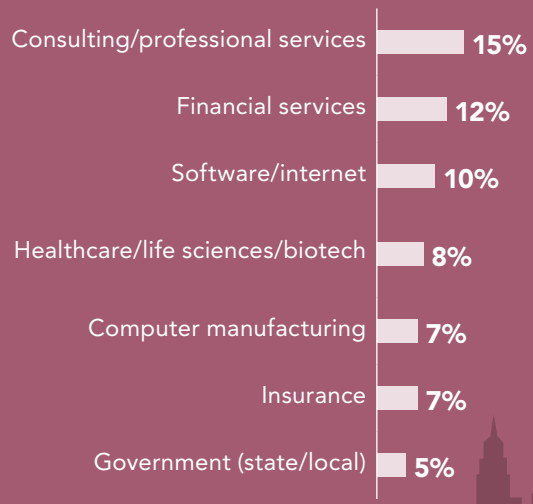
Industries vary considerably. Consulting/professional services were most represented at 15% of respondents, while financial services follows with 12%. Software/internet makes up 10% and healthcare/life sciences/biotech is 8% of the respondents. These are followed by computer manufacturing (7%), insurance (7%), and state and local government (5%). Secondary education, construction/architecture/engineering, manufacturing (non-computer), and retail/consumer products account for 4% each. Over half of respondents reside in the U.S. (56%), with Asia (14%), Europe (9%), Canada (6%), the Middle East (5%), and other regions (10%) following. Respondents come from enterprises of all sizes.

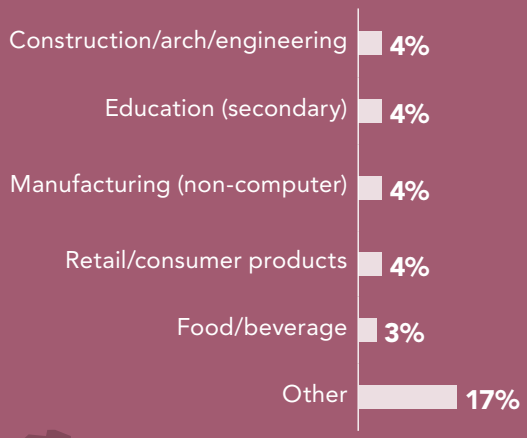
Position





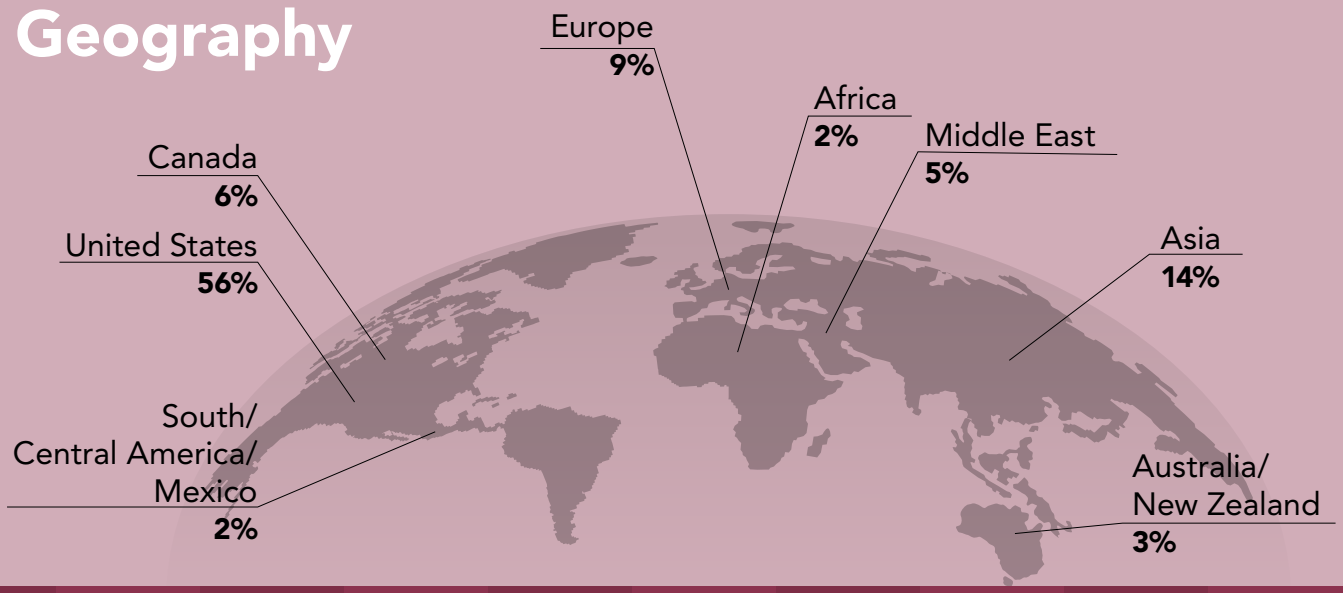
Industry



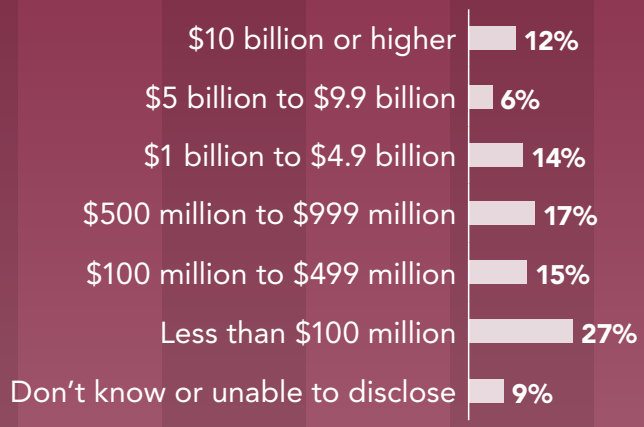


("Other" consists of multiple industries, each represented by 2% or less of respondents.)

Geography



Company Size by Revenue



Number of Employees

Based on the 181 complete respondents.

Executive Summary

Reducing speed to insight has long been an important objective, but many users are frustrated by delays in getting access to the right data and converting it into insights. Challenges are only growing as data gets faster, more varied, and more voluminous and organizations seek to deploy complex workloads that include artificial intelligence (AI) models and algorithms.

Decision-makers are at a disadvantage when they are forced to work with stale dashboards and analytics. Opportunities to use AI to personalize marketing, sales, and service at the point of engagement to increase customer loyalty are lost if data is old and irrelevant. Modern healthcare organizations depend on complete and accurate insights to improve patient care and understand population health. Financial services and insurance firms need timely insights to innovate with AI-driven products, evaluate risks, and detect fraud.

Faster insights can only be achieved if delays and bottlenecks throughout data life cycles are addressed using better practices and modern technologies. This TDWI Best Practices Report focuses on current challenges and how organizations can move forward to reduce time to insight and maximize the benefits of real-time data. Reducing time to insight involves a range of practices, data systems, and solutions to meet varied user and application requirements. The report discusses trends in AI-driven automation and augmentation, including generative AI.

Faster insights can only happen if delays throughout data life cycles are addressed using better practices and modern technologies.

Our research has found that poor data quality is a significant and common impediment. Seeing the relationship between data quality and data latency, organizations are putting a high priority on processes for improvement. They are also focusing on investment in data intelligence. This category of technologies and practices is critical to data quality and to streamlining data governance, which users often view as a hindrance to achieving data-driven goals. It includes data catalogs, metadata management integrated with data platforms, and master data management (MDM). These and related tools are vital to assembling knowledge that makes it easier to find relevant data, know about its quality, and track sensitive data for governance and security.

Organizations need to reduce bottlenecks and improve automation in developing and deploying data integration and preparation processes. Problems such as complex manual coding, redundancy, and excessive data movement are often major contributors to data latency. The research also identifies the problem of already numerous and growing data silos. Silos make it difficult to streamline data integration and ensure complete data quality and governance. The report discusses strategies such as data fabrics and data silo consolidation for solving data silo challenges.

Real-time data is growing as most machines, vehicles, systems, and equipment become connected data sources. Organizations want to tap the value of real-time data from sensors, mobile devices, telemetry tracking systems, and logs that record customer activity on websites. To ensure cost-effective value, data streaming and processing need to be part of a broader strategy for deploying the right technology to fit use cases and workloads. This report discusses the essentials for developing a successful strategy, including using data observability to gain a holistic perspective of how systems and processes are contributing to desired business outcomes.

Streamlining the Path to Data Insights

The ability to reach data insights faster is a competitive advantage. The less time it takes for executives, managers, and frontline personnel to access trusted and relevant data—and to develop and consume actionable predictive insights and AI-driven recommendations—the better they can respond to changing market trends, shifts in customer preferences, and operational challenges.

Data continues to explode in volume, variety, and velocity, with real-time streaming sources becoming critical to both human and automated decisions. Analysts, data scientists, and business users want to examine real-time data streams alongside historical structured data. Emerging open table format standards such as Apache Iceberg offer promise for reducing delays caused by having to work with different data systems for managing business transactions and analytics. Developers and data scientists want to use Iceberg to make it easier to use SQL to query accurate and consistent transactional data managed in large data lakes.

Interest is growing in unified data platforms that meld the best attributes of traditional data warehouses that hold structured data and data lakes that contain big volumes of semi- and unstructured data. Rather than confine users to limited extracts and aggregations, organizations are racing to harness the potential of all data for driving better business outcomes.

Additionally, distributed data solutions such as data fabrics and data virtualization can enable organizations to provide a single point of access to multiple sources without moving data. With the appropriate strategy for their environment, organizations are seeking to provide complete and continuous access to trusted data to fuel data-rich applications and enable development of shareable data products that can be offered to users through internal and external data marketplaces and exchanges.

As part of the ongoing digital transformation of business applications and processes, organizations are rapidly democratizing data and analytics to further data-driven decision-making. Users across lines of business (LOBs) and departments are pushing beyond the boundaries of legacy systems and platforms that do not scale as they seek agility and more complete, contextually rich insights. Technology trends toward AI augmentation of business intelligence (BI) and the development and deployment of large language models (LLMs) for generative AI are further accelerating and expanding users' capabilities.

Trends toward AI augmentation and LLMs for generative AI are further accelerating and expanding users' capabilities.

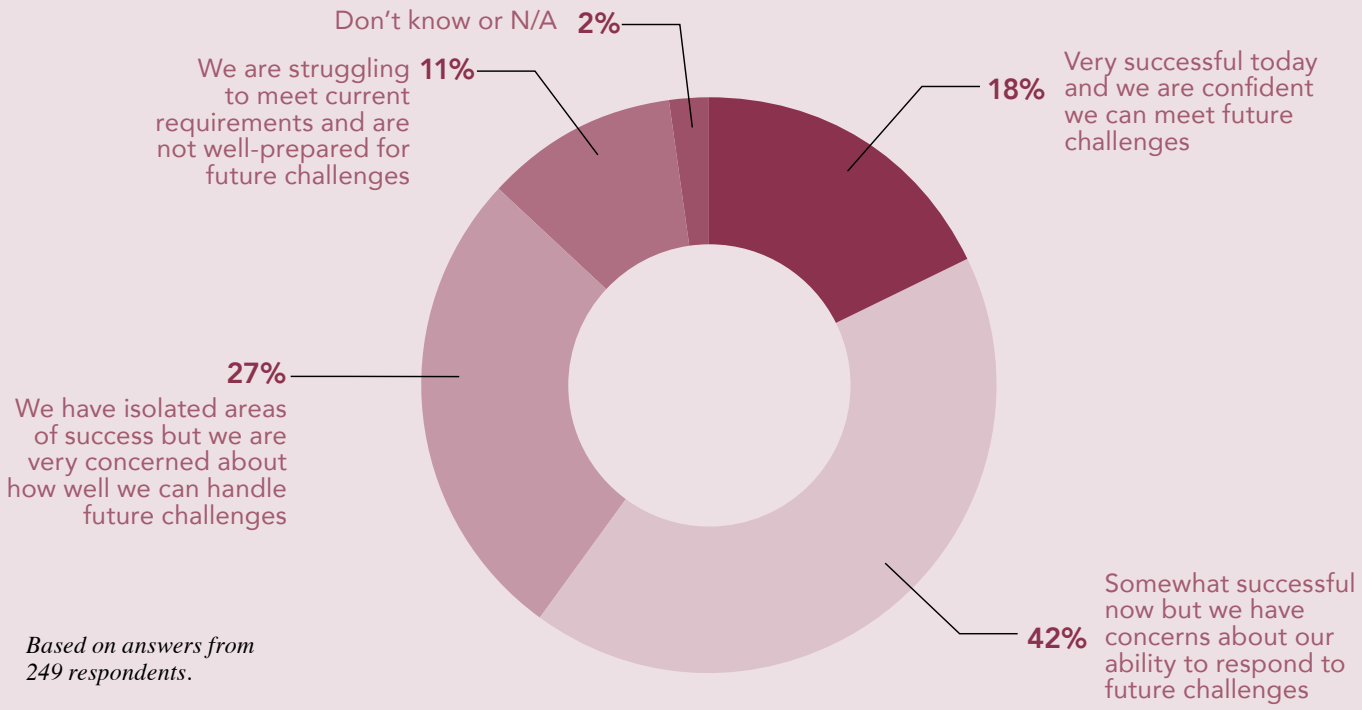
Enabling users to work with trusted, quality data sets and easily discover and integrate data is critical to reducing time to insight and improving productivity. However, as data democratization and deployment of self-service tools and AI-driven applications increase, organizations are under pressure to balance democratization with smarter, more agile data governance to ensure protection of sensitive data.

A major shared challenge is to expand data governance and protection to cover their entire environment of on-premises and cloud-native data systems. Thus, demand is rising for tools and data management systems that provide holistic perspectives for monitoring the effectiveness of data governance throughout environments. Organizations also need holistic data observability to determine whether data pipelines, connectivity, and transformation processes are meeting users' needs and are ultimately driving desired business outcomes.

This TDWI Best Practices Report will focus on identifying challenges organizations are facing in reducing time to insight and best practices for

Figure 1

Overall, how successful is your organization in using data systems to enable users to gain faster data insights? How confident is your organization that it can succeed as data and workload volumes, speed, and/or complexity increase?



Based on answers from 249 respondents.

overcoming them, as well as related technology trends. It will look at the role of real-time data streaming and how it fits with near-real-time data solutions so that organizations can develop a complete strategy for enabling faster data insights in the context of business requirements.

Improving speed to insight involves more than just accelerating the movement of data through pipelines. Faster data is not necessarily better data unless it is also sufficiently clean, accurate, valid, and complete enough to be fit for users' purposes. People and practice issues are as important as technology solutions. We will discuss both technologies and practices for eliminating obstacles to reducing time to insight. We will also examine organizations' current technology usage, satisfaction, and future deployment plans. We close the report with 10 best practices recommendations.

Overall Satisfaction and Business Benefits Sought

To begin, we asked about organizations' overall success in using data systems to enable users to gain faster data insights and their confidence that they can meet future data management challenges. Figure 1 shows that most respondents (42%) regard their organizations as somewhat successful now but have concerns about how well they will be able to respond to future challenges. Just over a quarter (27%) express greater concern about future challenges and say that they have only isolated areas of success now. With only 18% indicating that they are very successful today and confident in their organization's ability to meet future challenges, it's clear that most organizations are seeking improvement.

Defining the most important intended business benefits is critical for guiding technology investments. TDWI asked in which areas respondents believe their organizations could realize the most business benefits by increasing investment to reduce data latency, including by streamlining data processes and modernizing analytics to deliver faster insights.

Responses revealed that many organizations believe such investments would upgrade daily operational decisions (46%; figure not shown) and business and corporate performance management (46%), which is often used to guide operational decisions. Additionally, more than one-third of those surveyed (38%) say that investment would benefit financial management, including budgeting, planning, and forecasting. This suggests that the currency of historical data many organizations use for reports, dashboards, and spreadsheets for financial management is not adequate.

Daily operational and financial decisions suffer when based on inaccurate and out-of-date KPIs in BI dashboards. Performance management metrics based on old data and analytics also can lead to poor decisions that are out of step with current conditions. The resulting business mistakes can reduce trust in the data and contribute to inefficient and ineffective processes, budgeting, and resource allocation.

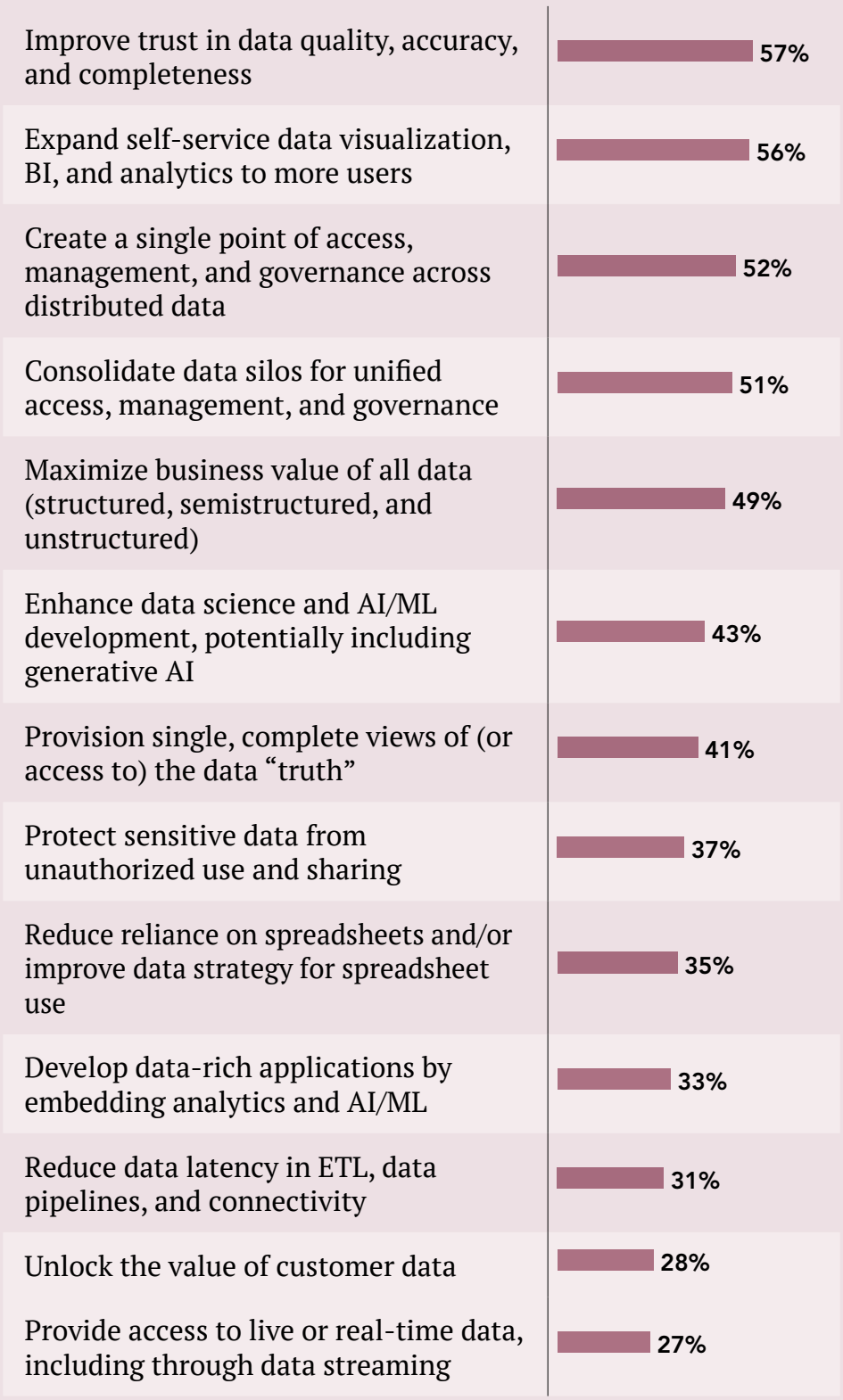
> *Decisions suffer when based on inaccurate and out-of-date key performance indicators in BI dashboards.*

Nearly as many respondents (43%) say that increased investment would enhance automation of decision-making in applications. This is an important area of focus today as organizations deploy online applications for real-time, personalized interaction with customers and

Figure 2

Which of the following are priorities driving modernization of your organization's data management and analytics systems and practices to achieve faster, more complete, and accurate data insights?

Based on answers from 230 respondents, who could select all that applied.



partners and need to supply data to embedded analytics and AI/ML models to drive those automated decisions. Over one-third of organizations surveyed (37%) say they would invest in reducing data latency specifically to support customer engagement and personalization.

Beyond decision automation, analytics- and AI-infused applications can empower people who work in customer-facing operations by giving them the most relevant insights for targeting sales strategies, marketing campaigns, and customer support. Nearly one-fifth of respondents (18%) say that increasing investment in reducing data latency would directly benefit marketing, sales, and service operations.

Modernization Priorities for Faster Data Insights

Turning our attention to modernization priorities for data management and analytics systems and practices, Figure 2 shows which are the most prevalent among survey respondents for achieving faster, more complete, and accurate data insights. Topping the list is improving trust in data quality, accuracy, and completeness (57%). Data quality steps that profile, cleanse, deduplicate, validate, and curate data are vital to delivering trusted data products that users regard as accurate and complete.

Nearly as many organizations (56%) prioritize expanding self-service data visualization, BI, and analytics to more users. Empowering all types of users with self-service functionality continues to be an important modernization driver. Users want to be free of IT dependence.

However, organizations need to ensure that self-service data consumption and analytics does not harm data quality, such as with data duplication problems that often arise due to easy establishment of siloed cloud-native data systems. These can become poor-quality data swamps that carry the risk of exposing sensitive data if not effectively governed. We can see in Figure 2 that about half of organizations (51%) surveyed are prioritizing consolidation of data silos for unified access, management, and governance.

At the same time, many respondents say their firms are prioritizing the creation of a single point of access, management, and governance across distributed data (52%). To accomplish this, organizations can deploy data fabrics and data virtualization layers that connect to and query multiple and distributed sources, aggregate the resulting data, and provide views of it from a single point of access.

> *52% of respondents say that creating a single point of access, management, and governance across distributed data is a modernization priority.*

To provision single, complete views of (or access to) the data "truth," which 41% say is a priority, organizations have technology options. For example, they could unify data in a single system or deploy a data fabric or virtualization layer, depending on which approach fits their use cases, data environments, and workloads. Some may choose to do both as part of a larger enterprise data strategy.

A significant percentage of organizations (43%) are also prioritizing enhancement of data science and AI/ML development, potentially including generative AI. Organizations are deploying AI models and algorithms across multiple

business operations, especially in customer-centric operations, to spot timely patterns in customer buying behavior and to deliver the right personalized engagement, product, or service offerings. TDWI finds that strengthening data management support for advanced analytics and AI initiatives is a major driver behind strategies for reducing data latency and time to insight. We will have more to say about the impact of AI in later sections.

Data Currency and Real-Time Data

Having the most current data is valuable for a range of operational and analytics use cases. Real-time data sources continue to grow in number and diversity to include machine data sensors, log files, online customer behavior records, geolocation systems, devices for telemetry data tracking, financial trading systems, social networks, in-game player activity, and video. Some sources produce continuous feeds of small-volume data and messages while others emit high-volume, multimegabyte bursts when triggered by an event or pattern of activity.

Data quality and consistency are leading data latency issues because doubts about whether the data can be trusted and the length of preparation processes to fix problems affect how quickly organizations can gain value from the most current data. From a quality perspective, some data is constant and stable; other data sources continuously change and produce data that varies in quality. Therefore, data quality processes for cleansing, validating, structuring, enriching, and completing the data cannot be one-size-fits-all. Organizations do not want these processes themselves to become unnecessary bottlenecks.

Data quality and consistency are leading data latency issues because they affect the value of having the most current data.

It is vital for users to know the level of data accuracy they are working with, particularly if they depend on near- or true real-time data feeds for actionable information in operational dashboards. Organizations need to monitor the collection of real-time data streams carefully for errors, missing data, and anomalies. Data curation and preparation processes must scale to handle not only the sheer velocity of real-time data but also the variety of formats, structures, and unique characteristics. Data collection, curation, and preparation must be efficient and resilient to avoid bottlenecks and downtime.

Real time is the sum of many parts. Real-time data processing involves an array of technologies designed to minimize latency and provide users and applications with data as soon after its creation as possible. This includes options such as fast networks, high-powered, massively parallel data processing, and in-memory computing where appropriate for meeting varied data workload requirements.

Because in many cases the data's value to decision-makers, analytics, and automated applications is associated with its age, time-stamping is important for recording the data's chronology. For some requirements, only the most current data is valuable. For example, for analytics looking for trends and patterns, chronological information is critical.

Some systems today embed AI/ML algorithms to automatically curate data and accelerate its preparation so it quickly achieves the most

desirable and readable output quality for visualizations such as dashboards, video or audio notifications and reports, and other carefully defined data products. Fast creation of virtual data warehouses to assemble processing resources and curated data on demand can shorten traditionally long, manual development. Data fabrics can offer virtual views of data, bypassing traditional ETL steps for landing data in a data warehouse.

The speed of inputting and curating the data must be matched by application functionality that enables users and automated decisioning systems to respond in the minimum amount of time. Today, continuous operational intelligence systems use real-time data and AI to detect physical situations such as maintenance or security issues, changes in online customer behavior, and deviations from expect patterns in systems performance. Applications can enable responses by taking automated actions or notifying responsible personnel.

Most organizations work with *near*-real-time data. What constitutes the highest level of data currency is often relative to the data available and user, analytics, and application requirements. TDWI asked research respondents what "real-time data" generally means in their organizations. The largest percentage of respondents (45%; figure not shown) say that data is available after a small but perceived and understood delay. The latency of this "near" real-time data could be seconds, minutes, or hours depending on the speed of processing and curation. Currency is set to fit the requirements.

> *Most organizations (45%) work with near-real-time data that is available after a small delay.*

Near-real-time can be a matter of user perception. A respondent who checked "other" noted that in their organization, "there is a mismatch in understanding between business and technical users; business users have an expectation of availability as soon as data is created whereas technical users anticipate a short delay." Another says that real-time data is simply "the data that is available at decision time."

The second-biggest percentage of respondents (35%) said their definition of real-time data is aligned with industry definitions: that is, data is available (including streamed) as soon as it is created and acquired. In these respondents' organizations, real-time data is available within a precisely defined time constraint typically within seconds or less (down to milliseconds).

Near- and true real-time processing contrast with traditional batch processing. Almost one-fifth of respondents (17%) say that users' most currently available data is provided when the batch processing job finishes. Data latency varies in these cases. For some batch jobs, latency could be minutes or hours while for others, it is from one day to multiple days, weeks, or months.

Traditionally, batch processing is most appropriate for handling updates to large volumes of data at designated intervals that are obviously much longer than required for near- or true real-time data. Queries are batched together instead of continuously run against smaller (sometimes "micro") batches of data or even just individual records. Regularly scheduled snapshots of sales, billing, and similar data for managed BI reports are often done in batches.

Organizations run complex analytics queries in batches for cost-effective use of resources. Real-time analytics processing is typically coordinated closely with tightly defined and

specific decisions, responses, and metrics that have urgency. Comprehensive data strategies should align workloads with processing options to deliver an appropriate level of latency and query performance.

Data Currency Types Available to Users

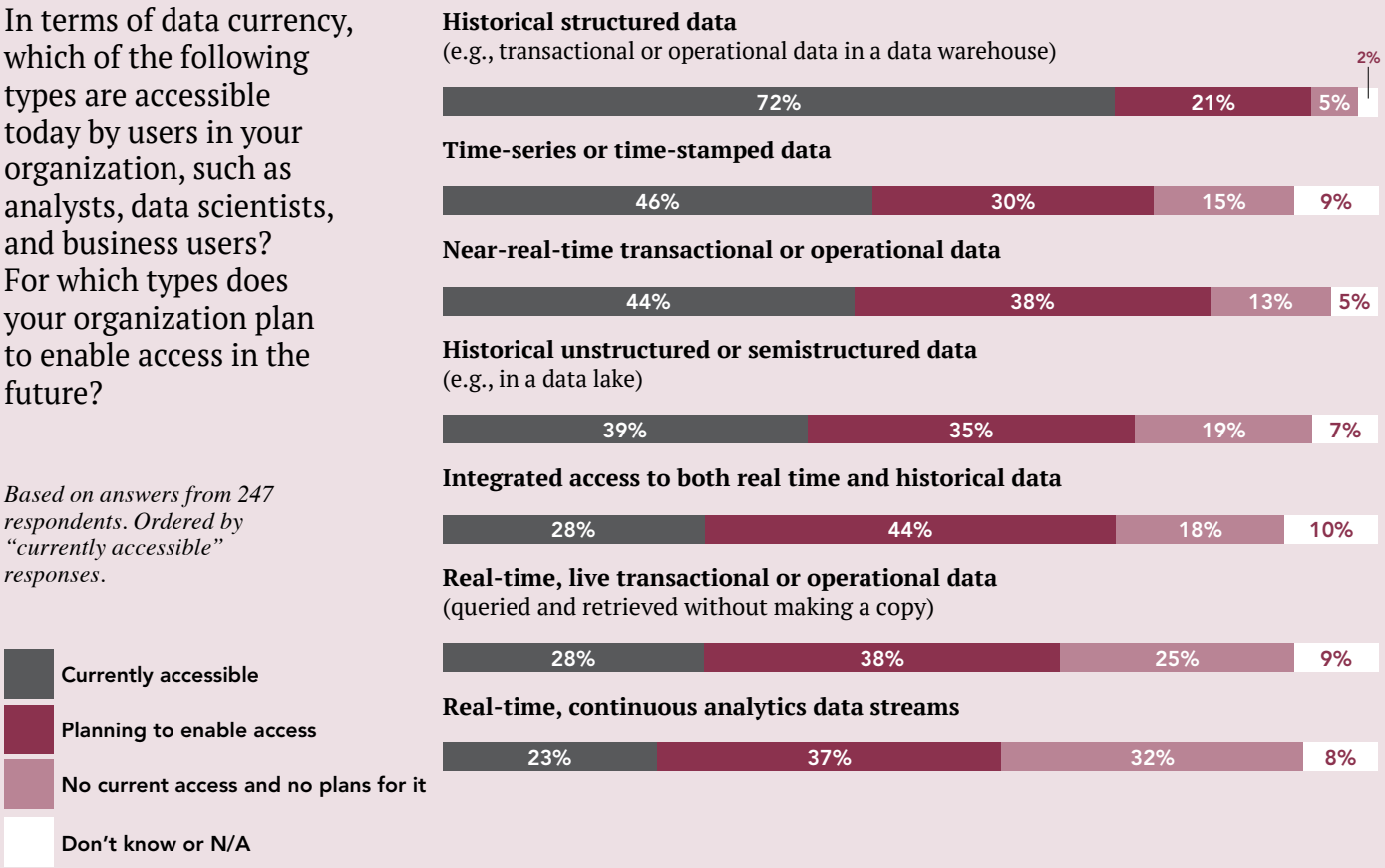
In terms of data currency, large organizations typically have a spectrum of data available to users, analytics, and operational applications. Figure 3 shows that among organizations of all sizes surveyed, not surprisingly, nearly three-quarters have available historical structured data such as transactional and operational data in a data warehouse (72% currently accessible; 21% planning to enable access).

Nearly two in five organizations (39%) enable access to historical unstructured or semistructured data, for example, in a data lake, while 35% are planning to enable such access. As an architecture optimized to capture a wide range of data types,

Figure 3

In terms of data currency, which of the following types are accessible today by users in your organization, such as analysts, data scientists, and business users? For which types does your organization plan to enable access in the future?

Based on answers from 247 respondents. Ordered by "currently accessible" responses.



a data lake is open to a variety of purposes. Primarily used for analytics and AI/ML, data lakes let each workload dictate the data preparation it needs. Data lakes can also be used to land and store transactional and streaming data.

At the other end of the spectrum, just over one-quarter (28%) have real-time, live transactional or operational data (that is, queried and retrieved without making a copy) available for access. Note, however, that a higher percentage of respondents (38%) plan to enable access to this type of data for driving operational use cases.

> *To support analytics, 37% plan to enable access to real-time, continuous analytics data streams.*

To support analytics, about the same percentage of organizations (37%) plan to enable access to real-time, continuous analytics data streams instead of waiting to analyze it in batches. These organizations often run predictive models and AI algorithms on real-time data sources such as telemetry and geospatial information from vehicles or mobile devices, online transactions, and customer behavior log files. Almost one-quarter (23%) say this data is currently accessible to users.

As discussed earlier, near-real-time data access is more prevalent than true real time. Figure 3 shows that 44% currently enable access to near-real-time transactional or operational data and 38% plan to enable access. To keep pace with fast data requirements, some organizations have developed data platforms to support real-time data warehouses that ingest continuous data streams or use fast-batch capabilities to frequently update selected near-real-time data.

As an alternative to long ETL processes to update data, some organizations use change data capture (CDC) technologies or event streaming platforms to continuously deliver incremental updates or notifications regarding changes to data and data structures. Organizations can also create data fabrics that use data virtualization and federation to connect to and query heterogeneous distributed sources; data fabric technologies aggregate the resulting data and provide views of it from a single point of access.

Finally, many data scientists and business users need integrated views and access to both real-time and historical data. Just over one-quarter (28%) say this is currently possible, and 44% plan to enable integrated access. The higher percentage for planning suggests growing interest in technologies and open standards for integrating access and views of real-time, transactional, and analytical data rather than forcing users to address each source separately.

Issues Hindering Timely Data Insights

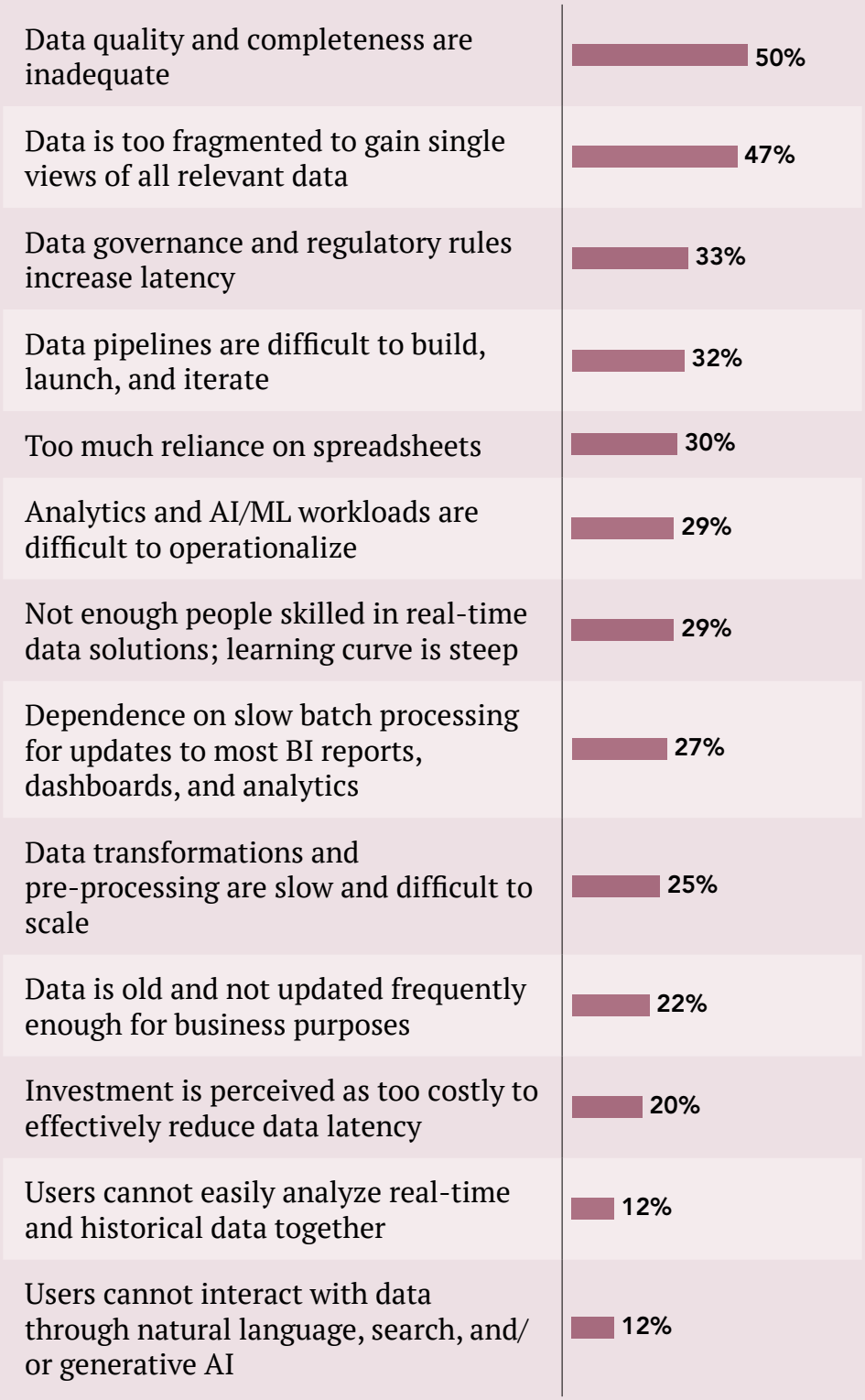
Multiple issues are to blame when users are unsatisfied with the data available for answering pressing business questions and solving problems. In Figure 4, we can see how hindrances rank in importance for organizations in our research. We will examine many of these challenges and potential solutions later in this report.

Reinforcing what we saw earlier, Figure 4 shows that inadequate data quality and completeness is the most common hindrance (50%). The importance of improving data quality, reducing data duplication, and increasing data completeness puts pressure on organizations to curate data

Figure 4

What are the biggest issues hindering your organization from enabling users to gain timely data insights for answering business questions and solving problems?

Based on answers from 245 respondents, who were asked to select at least their top five concerns.



effectively and monitor quality processes to ensure that data is fit for each user's purpose.

Organizations can use data intelligence tools such as data catalogs to know more about source data quality and use that knowledge to guide users to trusted and appropriate data sets.

Data currency is a key element of data quality. Nearly a quarter of organizations surveyed (22%) say that their data is old and not updated frequently enough for their business purpose. Slow data transformations and preprocessing, usually with ETL processes, often leave users with data that is too old; 25% cite this as one of their top hindrances. Slightly more respondents (27%) point to dependence on slow batch processing, which suggests that these organizations need to evaluate data platform solutions for near-real-time data access, real-time data warehousing, and real-time data streaming.

> *One-quarter of respondents say that slow data transformations and pre-processing is one of their top hindrances.*

Three additional issues contribute significantly to data latency:

- **Data silos.** Nearly half of respondents (47%) point to data being too fragmented to gain single views of all relevant data. Data silos full of inconsistent and often duplicate data make it difficult for users to gain a single view of the truth. Organizations need to evaluate strategies for consolidating data from silos into a single data platform or using a data virtualization layer to provide a single point of access to multiple data sources.
- **Data governance.** Data silos contribute to data governance and security concerns about exposing sensitive data. When organizations lack efficient processes for adherence to data governance rules and policies and are not taking advantage of automation, it often results in increased latency. One-third of respondents (33%) indicate that this is a hindrance.
- **Spreadsheets.** Spreadsheets are ubiquitous in most organizations for accessing data extracts and performing various functions for analyzing and visualizing data. However, spreadsheets can become poor-quality data dumps when analysts lack formal processes to incorporate new data and cleanse, enrich, and prepare it. Figure 4 shows that 30% of respondents regard "too much reliance" on spreadsheets as a hindrance.

Data Management Strategies for Reducing Latency

In this section, we focus on trends, challenges, opportunities to modernize data technologies, and practices for reducing time to insight. Modernization enables organizations to solve problems that contribute to data latency and to establish a foundation for growth. Orchestrating all the "moving parts" that must work together to deliver faster insights is never easy. It is important to have an overall data strategy that unifies decisions involving data technology and data process modernization. Data observability tools can help organizations gain holistic views of how well processes are fitting together.

For perspective on modernization priorities, we asked organizations which options are most

important right now for reducing delays and bottlenecks (see Figure 5). The top shared priority is deploying a data catalog, MDM, and/or data lineage tool to improve data intelligence (51%).

> *Metadata management such as with a data catalog is the foundation of data intelligence, which also includes MDM, data lineage tools, and more.*

Data intelligence is valuable for accelerating time to insight. Data intelligence includes technologies for obtaining and managing knowledge about diverse data. Foundational to data intelligence are metadata resources contained in a repository such as a data catalog or metadata management inside a data platform. The data intelligence category also includes MDM, data lineage tools, knowledge graphs, and semantic data integration. These tools and systems help organizations manage and coordinate data profiles based on master definitions as well as information about the data's origins, ownership, location, calculations, transformations, and documentation of taxonomies and classifications.

Also critical is streamlining and automating data governance; Figure 5 shows that 48% of respondents regard this as a priority. As noted earlier, users tend to view "defensive" data governance processes for protecting sensitive data and preventing its unauthorized use and sharing as impediments to developing and applying data insights. Using modern technologies and processes to streamline and automate data governance is therefore critical to making it less of a hindrance. Some organizations (24%) are setting up or participating in data marketplaces or exchanges to make it easier for users to locate, share, and use trusted data sets.

Organizations prioritize data consolidation and/or enabling virtualized data access. Consolidation of data silos is a priority for 46% of organizations surveyed. Nearly the same percentage of respondents (43%) are strongly interested in data warehouse and lake unification on one data platform.

Consolidating data silos and moving beyond traditional two-tier strategies by unifying data warehouse and data lake systems onto one data platform are significant trends in our research. Organizations have been taking these steps primarily to make analytics and AI/ML development easier. However, with the emergence of open table formats such as Apache Iceberg, organizations can develop strategies for bringing transactional and analytical databases together on unified data platforms.

> *Consolidating silos and moving beyond two-tier strategies by unifying data warehouse and data lake systems are significant trends.*

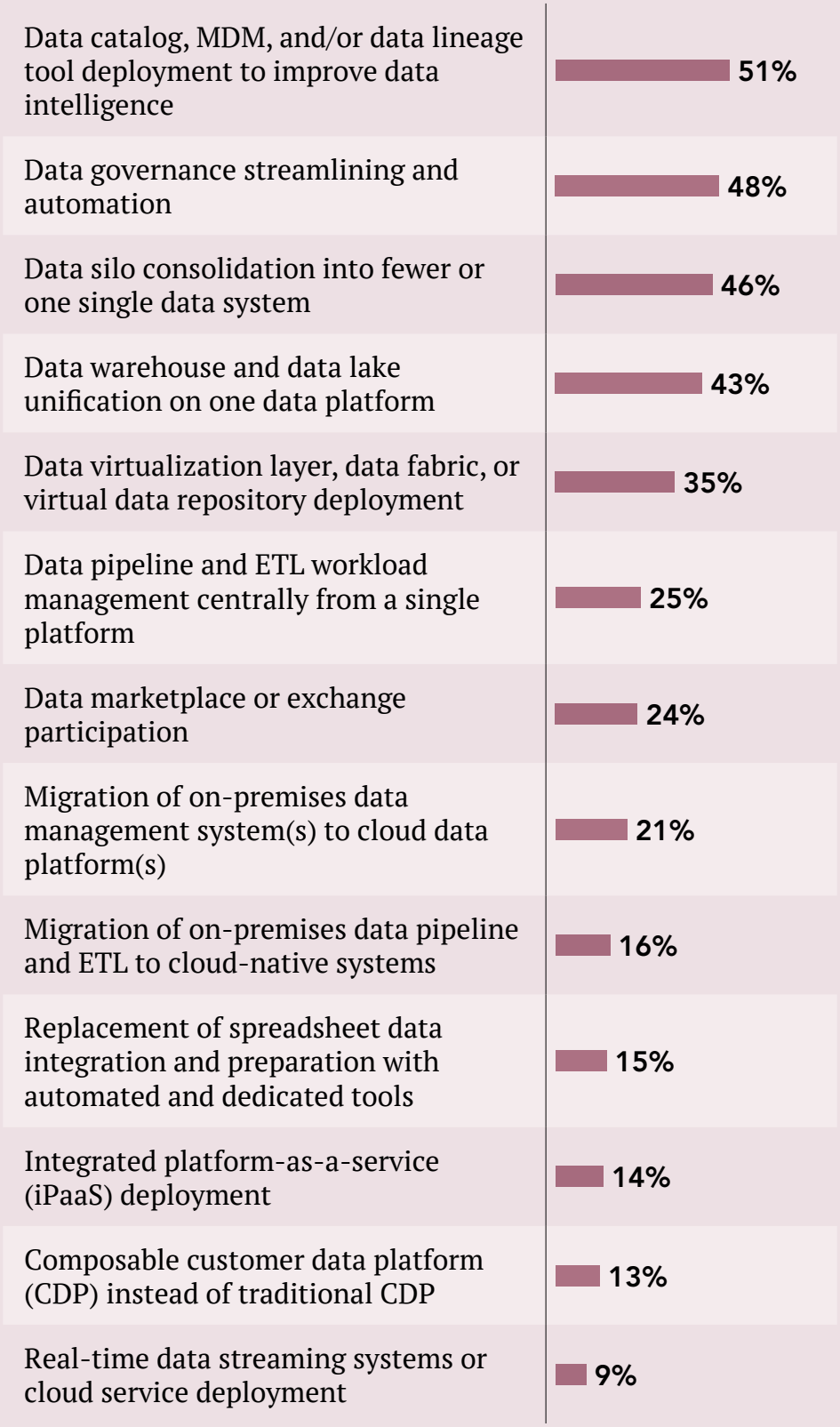
Alternatively, more than one-third of organizations (35%) prioritize adoption of a data virtualization layer, a data fabric, or a virtual data repository. The intention with these systems is to provide an abstraction layer that shields users from having to know multiple source data formats, definitions, and structural implementations. It also relieves organizations of having to move volumes of data into physical intermediate staging areas and data warehouses.

For integrating customer data, an emerging trend is adoption of a composable customer data platform (CDP); 13% are prioritizing this strategy. Traditionally, organizations that have a centralized CDP will direct users to access

Figure 5

Which of the following technology options are current priorities for your organization to reduce delays and bottlenecks that impede faster data insights, better business decisions, and turning insights into action?

Based on answers from 220 respondents, who could select all that applied.



this source. However, large organizations often continue to collect and store customer data in multiple systems. Mergers, acquisitions, and business reorganizations also create new customer data files that are not integrated into the central CDP. Customer data therefore exists in distributed data marts, data warehouses, data lakes, and application-specific databases such as CRM, sales management, social media, and e-commerce systems.

A composable CDP offers a virtual layer activating on-demand data access, data curation, and analytics. It can bridge distributed customer data sources. When complete customer data is available in a centralized resource, this agility can be critical to meeting diverse, ad hoc requirements that frequently arise in customer-centric operations such as marketing and sales, where users need different data to analyze rapidly changing trends.

Addressing Data Integration and Preparation Challenges

Interest is strong for solutions that can reduce the total time data engineers, data scientists, analysts, and business users spend on complex data activities such as building data pipelines and connectivity and preparing data through cleansing and curation. We asked organizations what percentage of their users' time is spent on these data tasks compared to time they are able to spend on value generation through analytics, including AI/ML. Half of respondents (50%) say that users spend 61% or more of their time on data integration, pipeline development, and preparation, with 16% indicating that users spend 81% or more of their time on these tasks (figure not shown).

The survey results are consistent with what we found when we asked this question in our 2020 report.[1] The lack of improvement suggests that as organizations add more users, they continue to struggle to reduce the time users spend on complicated data tasks.

Figure 6 offers an in-depth view of which data integration challenges organizations find most vexing. Governing and securing data as it is moved or copied tops the list with 69% regarding this as challenging; 29% say it is very challenging. Organizations need to monitor exposure risks while data is in motion during data integration or cloud migration. Organizations should evaluate data integration tools and data platforms that offer automated capabilities for monitoring data governance during loading, movement, and replication.

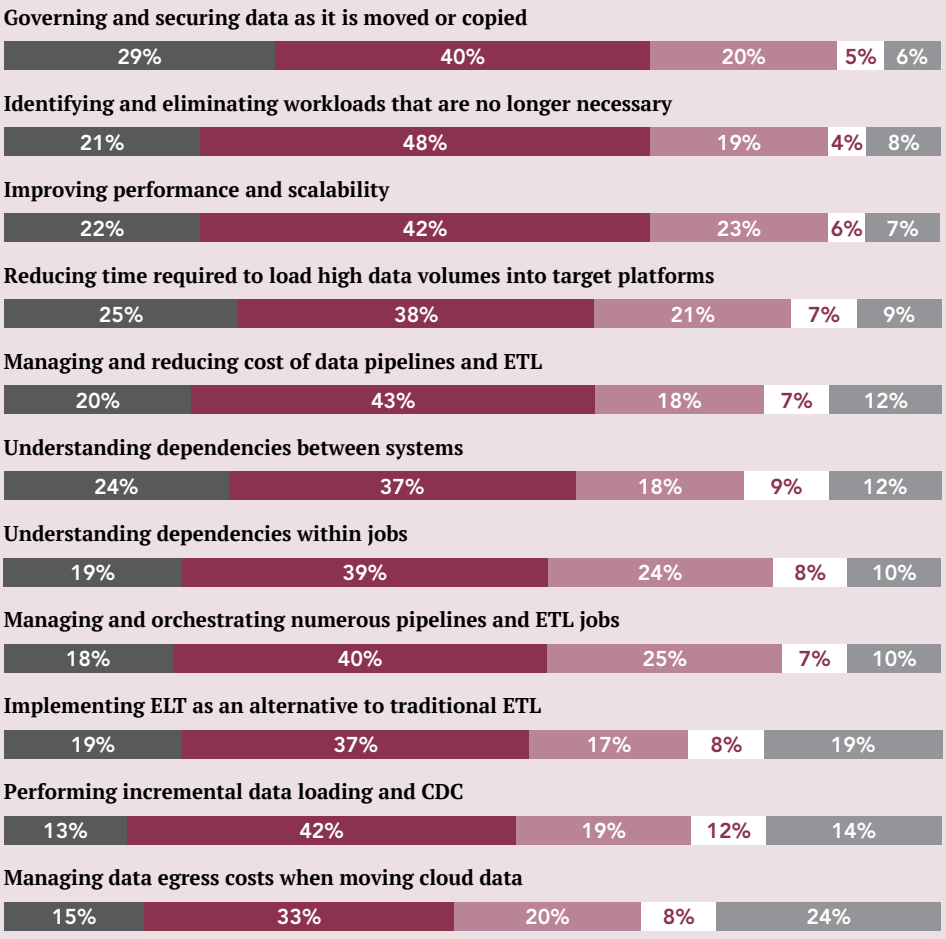
The same percentage of respondents (69%) indicate that identifying and eliminating workloads that are no longer necessary is a challenging task (21% say it is very challenging). Redundant and conflicting processes overtax computing and management resources.

Other challenges with high percentages suggest that organizations are having difficulty with holistic management of data integration and transformation processes. For example, 63% find managing and reducing the cost of data pipelines and ETL challenging, with 20% saying this is very challenging. More than half of organizations (58%) have difficulty managing and orchestrating numerous pipelines and ETL jobs (18% indicate it is very challenging).

[1] See 2020 *TDWI Best Practices Report: Faster Insights from Faster Data*, Page 21, Figure 8, online at tdwi.org/bpreports.

Figure 6

How challenging is it to address the following issues with your organization's data pipelines, data connectivity, and ETL or ELT processes?



Based on answers from 183 respondents. Ordered by combined "very challenging" and "somewhat challenging" responses.

- Very challenging
- Somewhat challenging
- Not too challenging
- Not at all challenging
- Don't know or N/A

Expanding Monitoring to Holistic Data Observability

Holistic data observability could help organizations troubleshoot problems and increase efficiency. Modern tools and DataOps practices for holistic monitoring and observability can improve how organizations identify and root out unnecessary and costly data pipelines and ETL jobs. Holistic observability enables better orchestration of data pipelines and ETL jobs as they grow in number and tap larger data volumes.

In Figure 5, we saw that one-quarter of organizations surveyed (25%) prioritize managing data pipelines and ETL workloads from a single platform. A unified, cloud-based data management and integration platform can serve as the focal point for modernizing tools and pipelines and making improvements for better processing scalability, faster connectivity, and efficient integration with data storage.

Holistic data observability would also help organizations understand dependencies between systems and between jobs, which Figure 6 shows

are significant challenges. If parts of a job fail, dependencies can lead to errors and missing data in other jobs, not to mention downstream in BI and analytics applications. Within jobs, if one process does not finish before a dependent process begins, it can slow down the entire process, causing increased data latency as well as potential for data errors. Figure 6 also shows that performance, scalability, and the time required to load high-volume data into target platforms are major challenges that demand data management tool and platform modernization.

> *Only 8% are very satisfied with end-to-end monitoring and observability; 30% are not fully satisfied but are making progress.*

In our research, only 8% of organizations surveyed are very satisfied with end-to-end monitoring and observability and their ability to solve problems; 30% say they are making progress toward goals but are not fully satisfied (figure not shown). The largest percentage of respondents (32%) report that they can monitor and address problems in individual systems, but end-to-end monitoring and observability are inadequate. The rest are either not satisfied or are just getting started with monitoring and observability and don't know yet (30%).

Siloed performance and availability monitoring is common in organizations. Our research indicates that organizations understand they need to advance beyond current monitoring to gain holistic management perspectives that will help them solve data integration and preparation issues that are causing unnecessary latency. Holistic monitoring can also be an important component of data fabrics to troubleshoot problems with viewing distributed data, data governance, and subpar data pipeline performance.

Importance of Data Intelligence and Data Catalogs

Data intelligence systems such as data catalogs can enable users to find trusted data sets, uncover data relationships, and learn more about data lineage, including how the data is currently being used and shared. Integrating metadata with master data and higher-level business definitions of customers, products, and other areas of interest reduces operational data confusion and accelerates analytics.

Figure 7 displays the diversity of organizations' current and planned use of different data intelligence systems and services. Topping the list is using a data catalog for a single data platform or application; 37% currently use and 37% plan to use one. A significant but somewhat smaller percentage of respondents (30%) say they are currently using a data catalog for a single department or project team; more plan to use one (41%). One-fifth of organizations surveyed (20%) currently manage metadata within a larger application, such as an ERP system (40% plan to use this approach).

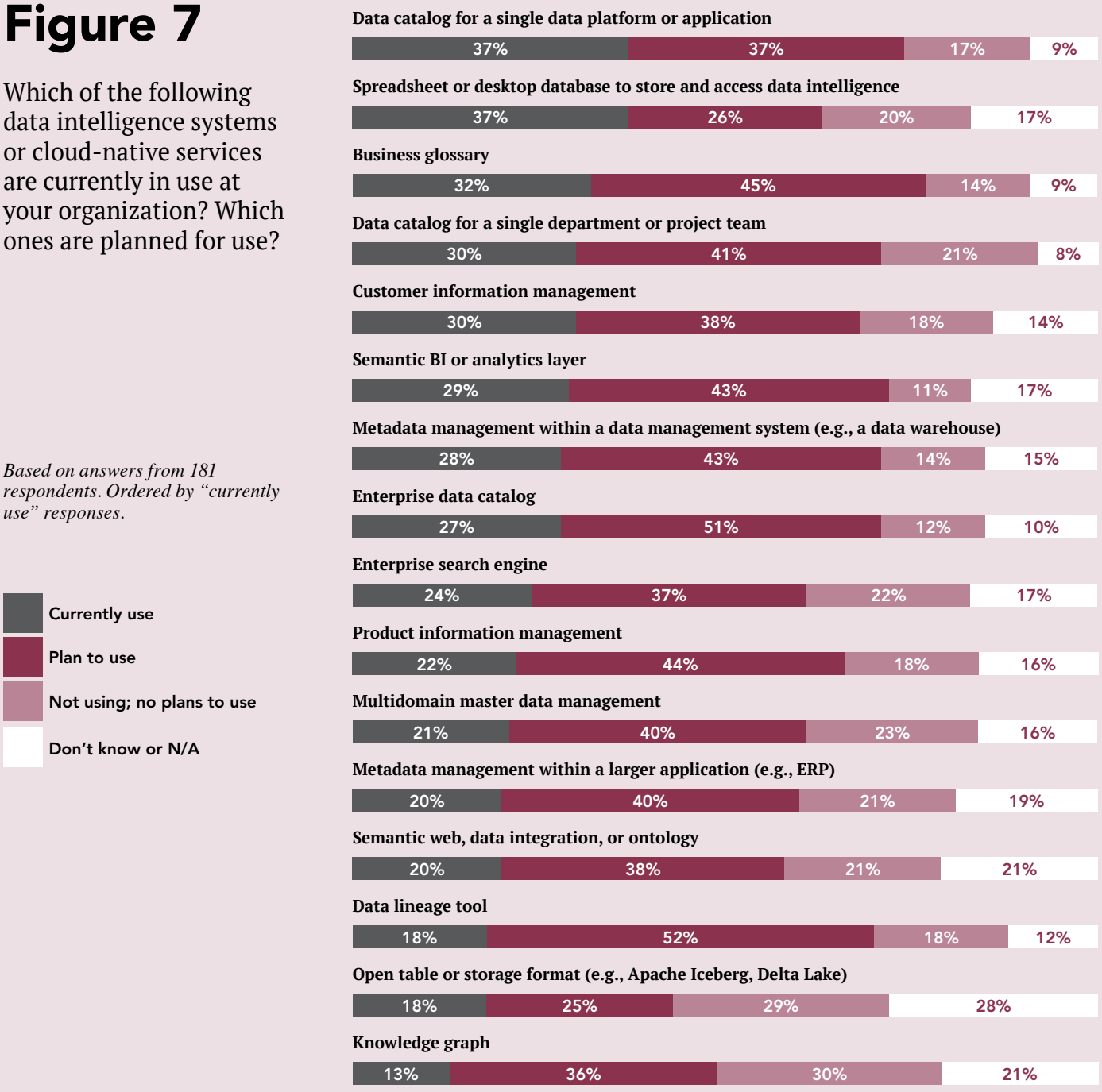
More than one-quarter (28%) have metadata management within a data management system (e.g., a data warehouse) and 43% plan to use this approach. Today, data management systems that provide metadata management capabilities are frequently cloud data platforms or data clouds.

Our research shows that fewer organizations (27%) currently use an enterprise data catalog, but significantly, just over half (51%) say they plan to use one. An enterprise data catalog offers

Figure 7

Which of the following data intelligence systems or cloud-native services are currently in use at your organization? Which ones are planned for use?

Based on answers from 181 respondents. Ordered by "currently use" responses.



advantages because of its scalability, scope in integrating metadata from multiple systems and applications, and its primacy as the single system of record for metadata and related information about data sets.

However, the more common data catalogs that are dedicated to single platforms, applications, data platforms, departments, and project teams still offer important value. One drawback of these types is that organizations may have inconsistencies between different single-purpose metadata management systems. Without an enterprise

system, users often find it difficult to integrate metadata and use it to locate data across systems to develop single views of objects of interest such as customers.

> *Without an enterprise data catalog users often face challenges in locating data across systems to develop single views.*

Almost two-thirds currently use or plan to use multidomain MDM. Like the metadata management variations, significant percentages of organizations surveyed are using or plan to use single-domain MDM systems for customer information management (CIM) and product information management (PIM)

Somewhat fewer (21%) are currently using multidomain MDM, but almost twice as many plan to use it (40%). Multidomain MDM integrates single master data domains such as customer, product, supplier, location, and asset. Multidomain MDM offers many advantages such as enabling users to gain insights into connections between customer preferences and product and supply data trends. Users can access integrated views of supplier and product master data rather than having to move from one separate, single-domain system to another and integrate information themselves.

Data Intelligence Modernization Priorities

With the research results of current and planned usage of data intelligence systems in mind, we now look at which modernization objectives are the highest priority in the next 12 months. The objective selected by most organizations in our research is to make it easier for users to search for, find, and collaborate on data assets (48%; see Figure 8). This highlights one of the most important potential benefits of a data catalog: enabling users at the point of data use to discover relevant, available, and trusted data. One-quarter of respondents (25%) prioritize lifting users out of metadata silos by unifying metadata and semantics from multiple data catalogs, BI tools, and data system instances. This usually involves setting up an enterprise data catalog.

Just over one-third (34%) say that an important objective is allowing users to add their own information and annotations to data catalogs to create richer data intelligence about data assets. On the cutting edge, 22% want to see AI-infused automation drive user recommendations about data assets. Generative AI capabilities will further revolutionize user interaction with data catalogs and other data intelligence systems by enabling users to ask natural language questions and work with chatbots that provide recommendations. This has the potential to accelerate data discovery and access to fuller information about data sets.

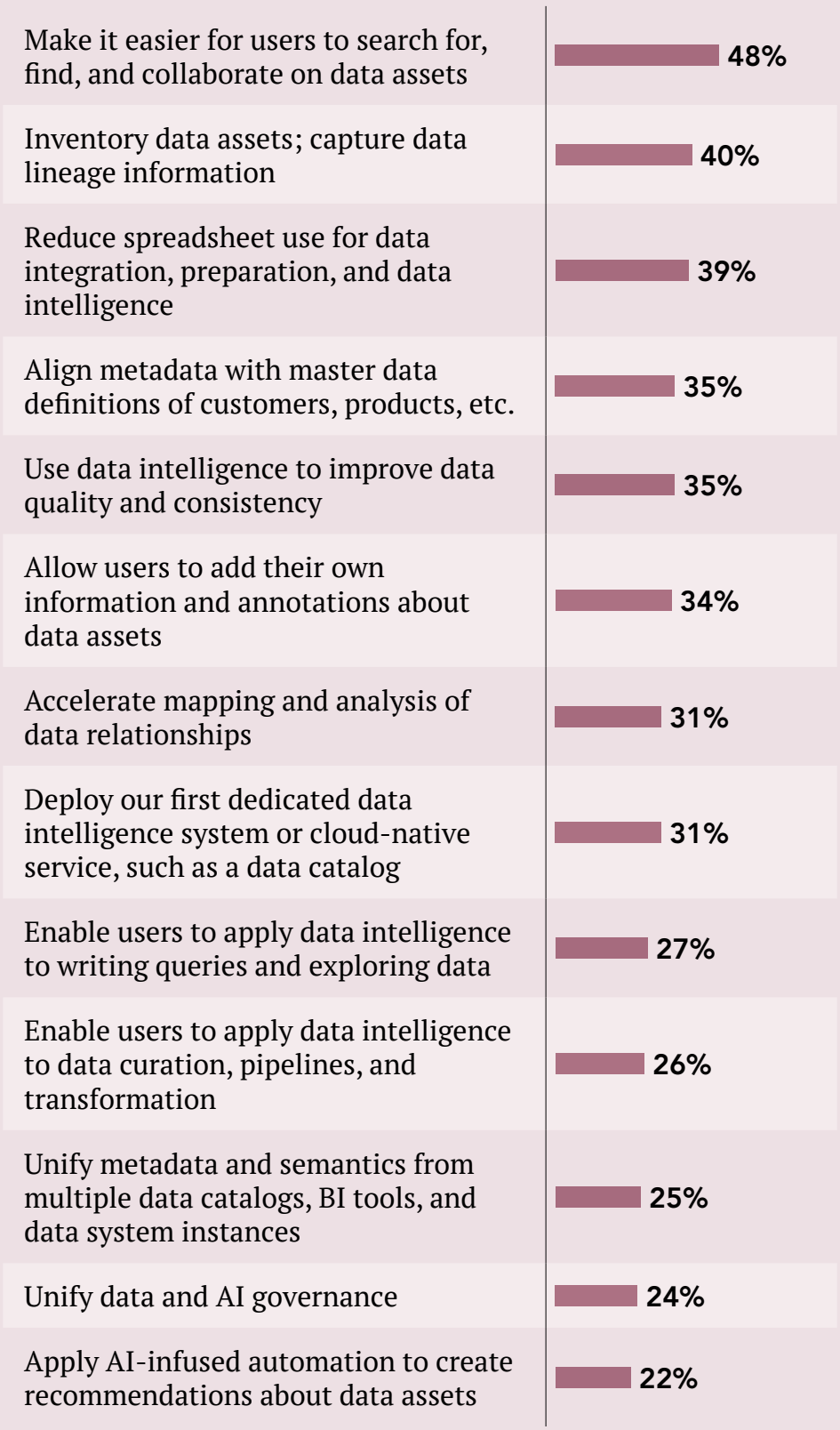
Improving how organizations develop inventories of data assets and capture data lineage information is a priority for 40% of respondents. These objectives are central to data governance. In Figure 7, just 18% of organizations surveyed reported currently using a dedicated data lineage tool; however, over half (52%) said they plan to use one. Data lineage capabilities are often included in data catalog solutions or in the toolsets of enterprise data platforms and data fabrics.

> *Reducing spreadsheet use for data integration, preparation, and data intelligence is a priority for 39% of organizations.*

Figure 8

To reduce data latency and increase speed to insight, which of the following objectives are the highest priority for your organization to achieve in the next 12 months with data intelligence systems and services, including a data catalog?

Based on answers from 178 respondents, who were asked to select at least their top five objectives.



The third-highest priority is reducing spreadsheet use for data integration, preparation, and data intelligence (39%). Single, complete views of data are important for BI reporting and analytics. Spreadsheets often make gaining complete data views difficult because they are single-user or departmental data silos. Often significant time is spent manually collecting, cleansing, completing, validating, and enriching spreadsheet-managed data, preventing analysts from gaining complete views of critical issues—for example, sales and marketing performance drawn from data generated in multiple channels.

Research shows three additional modernization priorities:

- **Aligning metadata with master data definitions of customer, products, and other topics (35%).** This is an important objective for integrating data intelligence. Data intelligence alignment is helpful in meeting another key customer analytics priority: accelerating mapping and analysis of data relationships (31% prioritize this). Some organizations are using knowledge graphs for this purpose.

- **Using data intelligence to improve data quality and consistency (35%).** Organizations surveyed regard poor data quality as perhaps the chief impediment to gaining faster speed to insight. Data intelligence is important to discovering and remedying data quality issues and enabling users to be informed about the quality of selected data sets.

- **Deploying an organization's first dedicated data intelligence system, such as a data catalog (31%).** With technology trends moving fast, it's easy to forget that many organizations are just getting started with data intelligence systems such as data catalogs. Nearly one-third of organizations surveyed prioritize taking this step in the next 12 months. Evolving AI and automation capabilities in data intelligence will enable organizations to avoid traditional issues such as manual development and updating and having to use interfaces that only technical experts understand.

AI-Driven Automation and Augmentation

AI is rapidly becoming part of everything in our world and an integral element of the technologies discussed in this report. The AI augmentation trend is central to driving automation for improving speed, accuracy, completeness, and scale.

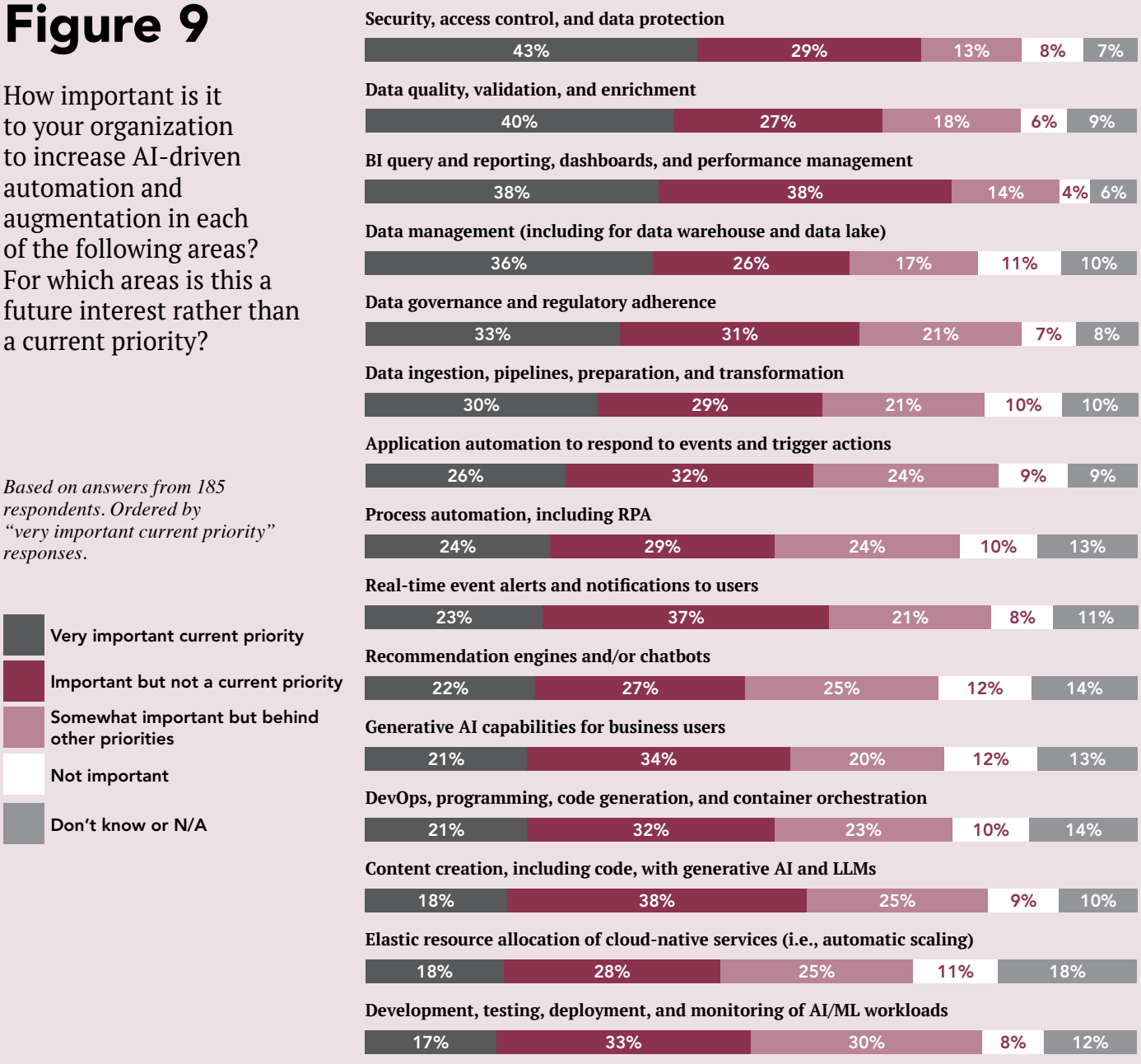
We asked respondents about the importance of AI-driven automation and augmentation to their organizations in several key areas. Figure 9 shows that for most organizations surveyed (43%), security, access control, and data protection collectively form their top current priority. Embedded AI models can help organizations monitor activity and identify unauthorized access quickly and accurately. Organizations can use models to analyze patterns for potential cybersecurity risks and identify threats in real time across many access points.

Organizations put nearly as high a priority on using AI to improve data quality, validation, and enrichment; 40% say this is very important currently and 27% regard it as important though not a current priority. AI-augmented tools or functionality embedded in data catalogs and data platforms can automatically detect data quality

Figure 9

How important is it to your organization to increase AI-driven automation and augmentation in each of the following areas? For which areas is this a future interest rather than a current priority?

Based on answers from 185 respondents. Ordered by "very important current priority" responses.



issues to make it easier to discover problems and remedy them appropriately. Organizations today are additionally prioritizing AI-infused capabilities for data governance and regulatory adherence (33%) and to improve data ingestion, pipelines, preparation, and transformation (30%).

> *40% of organizations want to focus AI-driven automation and augmentation on improving data quality, validation, and enrichment.*

About one-quarter (26%) say that their organizations are currently focused on using AI in application automation to respond to events and trigger actions. This increases the importance of streamlining data quality and curation. Automated decisions in data-driven applications require continuous access to trusted data for real-time response to input such as customer behavior in response to personalized offers or discovery of manufacturing and equipment maintenance problems. Just under one-quarter (23%) are prioritizing AI for real-time event alerts and notifications to users, with 37% indicating it is important but not a current concern.

Priorities for Improving User Experiences

Although not yet in widespread use according to our research, several items in Figure 9 will eventually have a major impact on users' data interaction. These include recommendation engines and/or chatbots (22% say these are a very important current priority), generative AI capabilities for business users (21%), and content creation (including code) with generative AI and LLMs (18%). All show more respondents regarding these as "important but not a current priority." However, they are likely to become more important in the future as experience grows, technologies mature, and organizations modernize data governance and data management practices to adapt to their use.

To conclude our research, we asked organizations which initiatives are top priorities overall for reducing users' time to insight. Not surprisingly, given our research thus far, the largest percentage of respondents (48%; see Figure 10) say that improving trust in data quality, consistency, and completeness is one of their top five priorities. As noted, there is strong interest in applying AI-infused capabilities to data quality, which is underway with leading current solutions. Other key priorities include:

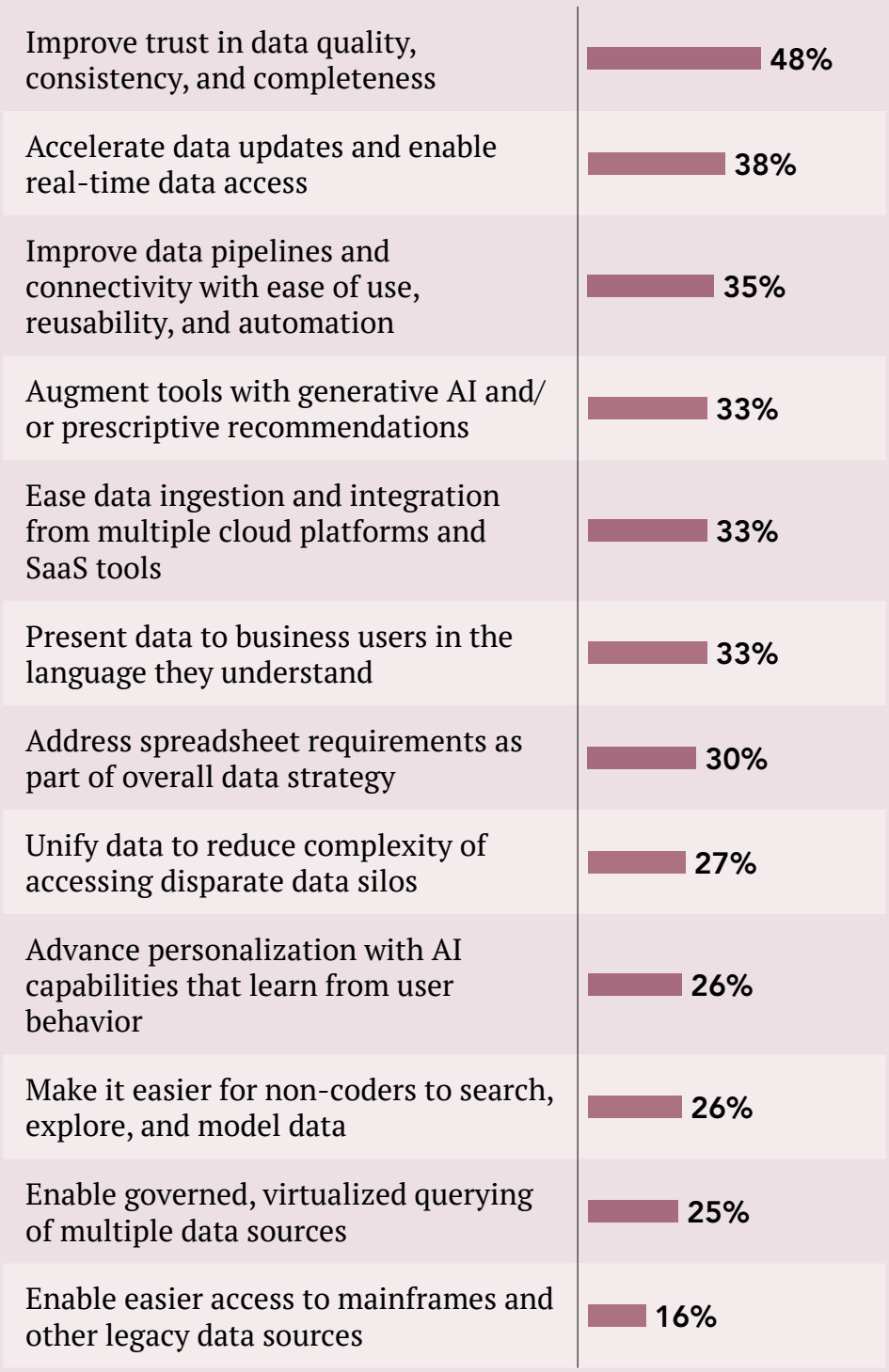
- **Accelerate data updates and enable real-time data access (38%).** The second-highest percentage of respondents share interest in reducing data latency toward real time. This aligns with our earlier discussion about the array of options available for accelerating updates and enabling real-time data access. Organizations need to interpret user requirements carefully to deploy appropriate solutions.

- **Improve data pipelines and connectivity with ease of use, reusability, and automation (35%).** This result indicates the level of interest in reducing unnecessary technical complexity and duplication (in data pipelines and connectivity) and applying automation where possible. Regarding ease of use, related results in Figure 10 show that 26% are interested in making it easier for non-coders to search, explore, and model data.

Figure 10

Which of the following objectives are top priorities to address in the next 12 months for your organization to improve user experiences with data and enable them to reduce time to insight?

Based on answers from 190 respondents, who were asked to select at least their top five concerns.



Automation plus low- and no-code solutions for data discovery, data pipeline development, and modeling are important to addressing these priorities.

- **Augment tools with generative AI and/or prescriptive recommendations (33%).** For a significant percentage of organizations surveyed, it is a priority to use AI to provide "copilot" functionality that can respond to requirements and shorten the path to contextual and actionable insights. Figure 10 shows that 33% find it important to present data to business users in the language that they understand and 26% say it is a priority to advance personalization with AI capabilities that learn from user behavior. Generative AI will expand how users can query data in their natural language.

Recommendations

Here are 10 best practices for reducing time to insight and maximizing the benefits of real-time data. Our research shows that these are key for meeting growing and evolving user requirements and taking advantage of faster and more scalable data processing, management, and integration technologies.

Address data quality weaknesses that increase latency. Organizations that aspire to be data-driven know that data quality is fundamental to operational decisions as well as advancing with analytics to understand customer behavior and set business and financial strategies. Data that is too old, incomplete, or inaccurate reduces trust, which increases business delays as users wait for IT to deliver better data or seek sources on their own. Organizations in our research place improvement in data quality among their highest priorities. Take advantage of modern, AI-infused data quality tools, data catalogs, and data platforms to remedy problems.

Align near- and true real-time data solution capabilities with requirements. Our research finds that most organizations work with near-real-time data. However, there is more than one way to reduce data latency. Multiple approaches can be cost-effective and relevant depending on requirements. Solutions include updating data incrementally, employing the power of modern data platform processing, and providing a virtual single point of access to real-time data sources. Organizations should develop an integrated data strategy that includes multiple approaches. Examine whether users and applications need near- or true real-time data access—or whether traditional batch processing can suffice.

Address data silo challenges through consolidation and/or data virtualization. Data silos are a common challenge in our research. As lines of business and departments take advantage of cloud computing, silos unfortunately can multiply. Consolidating data silos into single, unified data platforms is a strong trend. However, distributed data is likely to always remain part of modern data environments for a number of reasons, including data residency laws and the use of multiple cloud data storage platforms. Many organizations therefore need a comprehensive strategy that may include both data silo consolidation and using data virtualization to drive holistic data fabrics for broader access to distributed data. Develop a strategy that satisfies current and future user requirements with the appropriate technology for data management.

Organizations may need a comprehensive strategy that includes both data consolidation and a data fabric using data virtualization.

Apply modern technologies for effective but less intrusive data governance. Addressing data governance priorities for protecting sensitive data and guiding users to trusted data sets is a priority. However, a significant percentage of organizations surveyed acknowledge that users regard data governance as a hindrance to faster data insights. Modern systems and technologies apply AI-driven automation to streamline data governance. Organizations should evaluate solutions. They should also train data stewards, who can provide fellow users with expertise in data governance, regulatory compliance, and using data catalogs to discover trusted data sets.

Maximize the value of data catalogs, MDM, and additional data intelligence. The more knowledge organizations can gather about data and how it is defined and related, the more easily users can discover and integrate data they need. Metadata management, including using data catalogs, provides essential information such as data lineage. Data catalogs and metadata management need to be integrated with other data intelligence in MDM, semantic data integration, and knowledge graphs. Data intelligence is critical to data governance and enabling users to accelerate trusted data discovery across distributed data environments. Assess current capabilities and determine how modernization could maximize value.

Address delays caused by data integration, pipeline, and ETL challenges. Our research finds that users continue to spend most of their time on data integration and preparation, leaving less time for analytics to solve business problems. Data integration often involves significant manual coding and custom connectivity development. Organizations should modernize to take advantage of greater scalability, ease of use, and AI-infused automation. Organizations should also consider alternatives to traditional ETL. These include ELT to load raw data first into cloud data platforms and take advantage of powerful processing for transformation. Data fabric alternatives use data virtualization to significantly reduce the need for ETL and avoid slow and complex data movement. Ensuring easy access to data intelligence makes it easier to discover trusted data and avoid mistakes.

Enhance user empowerment with modern self-service capabilities. Data democratization and implementation of self-service technologies for BI, analytics, data integration, and data preparation remain strong trends. AI is advancing self-service by enabling natural language querying, driving smart automation, and augmenting user interaction with generative AI copilots such as chatbots that answer questions and provide recommendations. However, organizations should not stand pat with current self-service. Ensure that users can experiment with and ultimately maximize the latest capabilities. Update data management, governance, and integration to support new types of data interaction by a larger number and variety of users.

Reduce data latency that holds back analytics, including embedded in applications. A significant percentage of organizations in our research are prioritizing investment for reducing data latency in data science and AI development. Near- and true real-time analytics are vital for responding at the right time to customer and market trends, events such as fraud and cyberattacks, and problems in manufacturing, supply chains, and operational

processes. Data-driven applications often need continuous data updates to feed embedded analytics for automated decisions. Evaluate the range of technology solutions. These include faster and more scalable data platforms, data fabrics for eliminating complexity in accessing and governing distributed data, real-time data streaming, and fast batch processing.

> *Near- and true real-time analytics, including capabilities embedded in modern applications, are vital for responding to trends, events, and problems at the right time.*

Improve monitoring with holistic data observability. Data observability broadens standard monitoring beyond siloed views of technology systems and network performance. Observability focuses on visibility into the total health of data environments and the ultimate impact on the business of data-related delays and downtime. Our research finds that organizations need better visibility into end-to-end data life cycles. This can help enterprises prioritize modernization and identify opportunities for automation. With observability tools, often implemented along with DataOps methods, organizations can uncover problems such as disconnected and redundant data pipelines and ETL processes. Evaluate the potential of data observability.

Provide data literacy training and attend to people issues for a better data culture. Reducing time to data insights is not just about technology; it's about people. Organizations often fall short of realizing the value of technology investments due to lack of training and support for a healthy data culture. Data literacy training is important for raising individuals' proficiency in understanding what data means and their ability to communicate and share insights. Organizations can also use data literacy training to increase recognition of responsibility for data quality and governance. Supported by the right technologies and practices, data literacy can accelerate data insights and improve the data culture.

zoominfo

Achieve Go-to-Market Success with AI and Operations Excellence

Accurate, Accessible Data at Scale and Automated Data Management

Access a Trusted B2B Data Source

Data isn't just a part of your AI strategy—it is your AI strategy. Yet most businesses struggle to make their data accurate, reliable, and seamlessly integrated and accessible across systems and stakeholders. ZoomInfo reduces manual data prep, fuels analytics, optimizes AI models, and increases overall speed to insight with continuous, flexible access to trusted B2B data.

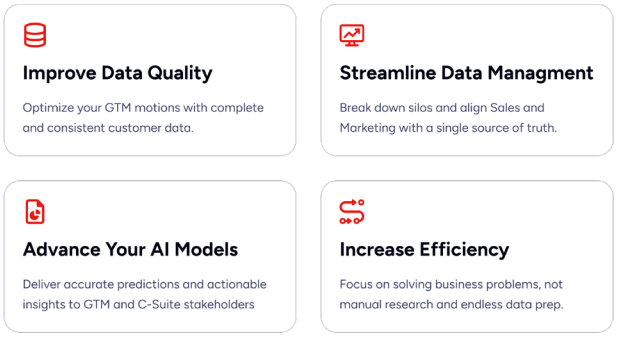
When it comes to data quality commitment, ZoomInfo:

- Updates over 1 billion data points weekly
- Employs a team of 150+ data professionals
- Invests more money in data annually than our five closest competitors combined

The result? The freshest, most actionable go-to-market data and insights.

Learn about ZoomInfo Data

Accelerate Insights with Seamless Data Discovery



TDWI found that the highest priority for companies is to make searching, finding, and collaborating on data assets easier.

ZoomInfo Data-as-a-Service (DaaS) partners with leading cloud providers such as Snowflake, Google Cloud, Databricks, and AWS to deliver fresh data to your cloud environment.

- Ingest data at scale for easy discovery
- Combine and enrich distributed data for a unified view
- Democratize data for generative AI, ML, and LLM use
- Collaborate with consistency

For complex use cases, ZoomInfo's data experts assist with strategy and implementation, and tailored consulting services.

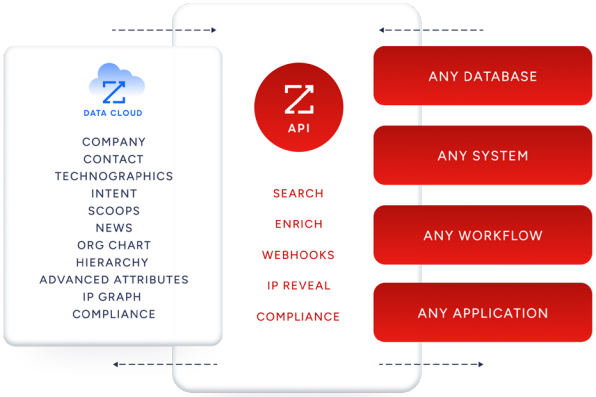
Learn about ZoomInfo Data-as-a-Service

Accelerate Insights with Real-Time Data Delivery

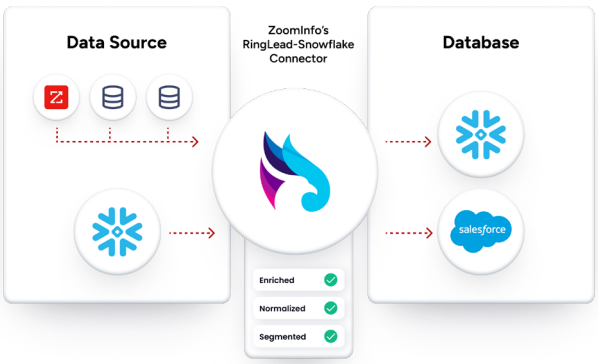
Data changes over time as people switch roles and companies evolve. Not staying current hinders decision-making, undermines data trust, and misses AI-driven outreach opportunities.

Flexible ZoomInfo APIs give customers access to the freshest data that integrates seamlessly into any business system. Webhooks push real-time data updates into those systems.

Decision-makers can anticipate risks, customer-facing teams can personalize their approach, and data analytics and models have more accurate outcomes.



Reduce Operations Latency with Automated Data Management



Faced with expanding data silos and tedious data prep, the majority of organizations TDWI surveyed expressed concern over how well they can respond to future data management challenges.

Too many businesses rely on spreadsheets and other manual methods that are time consuming, error prone, and obstruct single views of data.

ZoomInfo Operations is a no-code, automated data management engine that normalizes, dedupes, segments, and enriches data.

- Eliminate manual spreadsheets and redirect efforts to AI/ML projects
- Increase operational efficiency and trust with clean, complete data and automated enrichment
- Reduce IT dependence for data quality improvements
- Improve cloud governance and security
- Ensure continuous improvement

ZoomInfo's data management integration with Snowflake bridges the gap between first- and third-party data. Using Snowflake tables and ZoomInfo data, users build identical data sets in different systems without manual coding, redundant jobs, and excessive data movement.

Learn about ZoomInfo Operations

ZoomInfo (NASDAQ: ZI) is the trusted go-to-market platform for businesses to find, acquire, and grow customers. It delivers accurate, real-time data, insights, and technology to over 30,000 companies worldwide. ZoomInfo is a leader in data privacy, GDPR and CCPA compliance, and has earned numerous data security and privacy certifications.

TDWI Research provides research and advice for data professionals worldwide. TDWI Research focuses exclusively on data management and analytics issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of data management and analytics solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.

A Division of 1105 Media
6300 Canoga Avenue, Suite 1150
Woodland Hills, CA 91367

E info@tdwi.org

tdwi.org