AWS FOR DATA

CDO Agenda 2024: Navigating Data and Generative AI Frontiers

Thomas H. Davenport, Randy Bean, and Richard Wang

Table of contents

aws

Executive summary

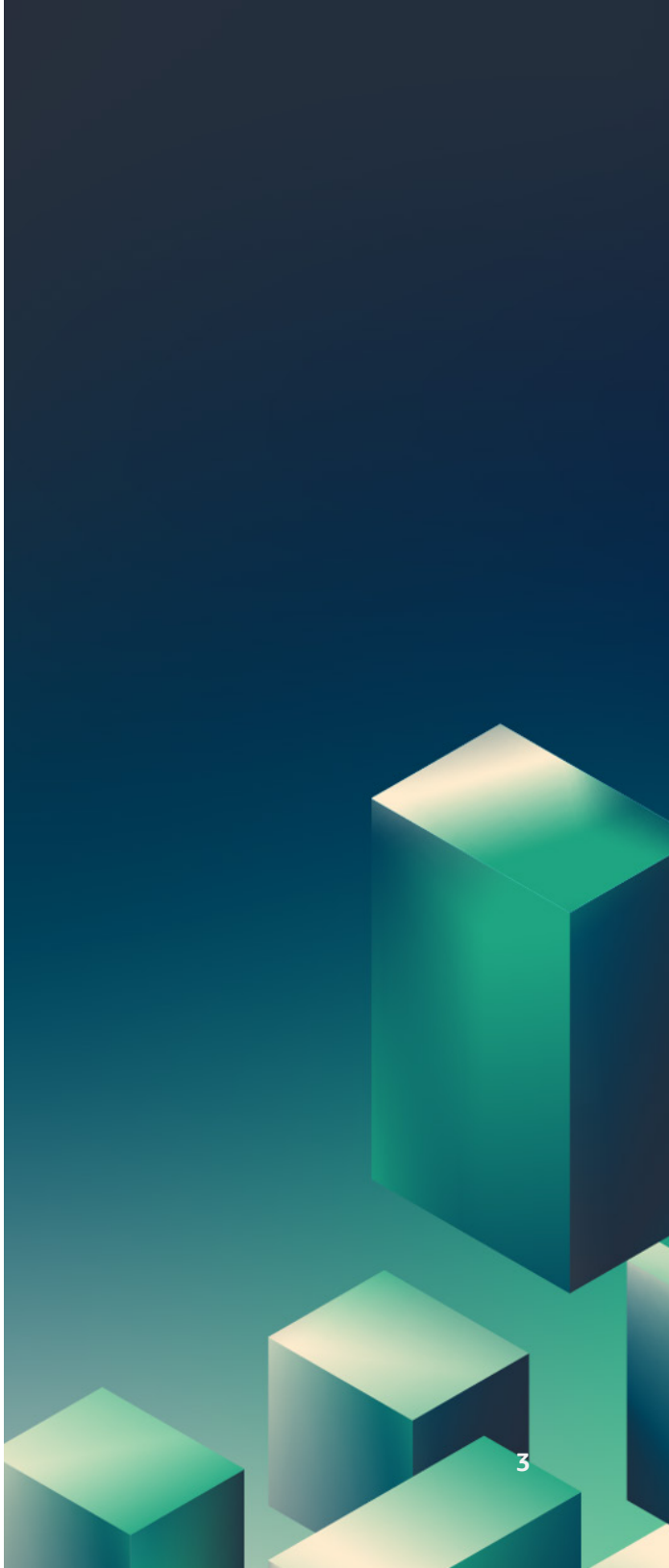
Collecting data, transforming it, and putting it to use is essential to remain competitive today. Virtually every organization is actively pursuing the goal of becoming more data-centric, enabling them to swiftly uncover and respond to valuable insights. With generative AI, the power and potential of organizational data is bigger than ever before. Organizations often appoint chief data officers (CDOs) to lead them toward data-driven success.

In the summer of 2023, we conducted a global study to understand how this role is evolving, what the key priorities and roadblocks are for CDOs, and how they tackle generative AI. This is the second annual report on CDOs, chief analytics officers, chief AI officers, or some combination of these titles or equivalent (hereafter described as CDOs). This is one of the largest CDO studies in the market, with a survey of 334 respondents with CDO or equivalent titles and level positions. It includes qualitative interviews with 12 leading CDOs.

One of the key findings from our study is that CDOs need to show a visible value for their efforts—in part by emphasizing analytics and AI. As a relatively new C-level role, CDOs have more topics than ever to focus on, which makes it difficult to standardize the job description. We found that CDOs also need to be flexible and agile to change their charter with changing technologies, such as generative AI.

CDOs are excited about the possibilities of generative AI, even though their companies are mostly experimenting with it. They feel that data and data strategy will be critical to success with generative AI, and they are in the early stages of that transformation. Other topics that are a top priority for CDOs include data governance and cultural change toward a data-driven organization. The savviest CDOs are prioritizing change management, communication, and evangelism, and they consider making other executives successful as critical to their success in the role.

This research report comprises key insights from our study and conversations with esteemed CDOs on how they're setting themselves up for success in the new generative AI era.



Key learnings

1 Creating visible value is still a key focus for many CDOs

Forty-four percent of CDOs still define success as achieving business objectives, as opposed to technical accomplishments (only 3 percent), and mark analytics and AI as keys to providing value. Other approaches include literacy training, councils across the organization, approaching data management use case by use case, and the data product management approach.

2 CDOs are enthusiastic about the potential of generative AI

Generative AI dominated the discussion. The majority of CDOs (80 percent) agreed that it would eventually transform their organization's business environment, but they don't want to abandon existing data-related initiatives in favor of generative AI yet.

3 The biggest roadblocks to generative AI—data quality and finding the right use cases

A large portion of CDOs (46 percent) rank data quality and finding the right use cases as the two biggest challenges for realizing the potential of generative AI, followed by (creating) guardrails around responsible AI, security, and privacy of data. Customer operation, such as customer support and chatbots, was the top generative AI use case (44 percent), followed by overall personal productivity (40 percent), and software code generation (36 percent).

4 Data strategy is crucial to generative AI success

More than half (57 percent) had not yet made necessary changes to their company's data strategies to support generative AI, but a majority (93 percent) of CDOs agreed that data strategy is crucial for getting value out of generative AI. A quarter of CDOs are pursuing data integration and cleaning, while nearly one-fifth are surveying data to understand what might support generative AI use cases.

5 Culture comes first, but challenges persist

Culture is increasingly critical to effective data use, but CDOs are taking a "slow and steady" approach to changing it. Culture initiatives are a major focus for over half of the CDOs surveyed, including data literacy programs and change management.

6 CDOs embrace new data governance strategies

Similar to last year, CDOs spend a substantial part of their time focusing on data governance activities (63 percent in 2023 vs. 44 percent in 2022). The new methods of establishing governance include an "enablement" focus—making it easier to do the right thing with data—and common data platforms.

7 Data-driven transformation is a team sport

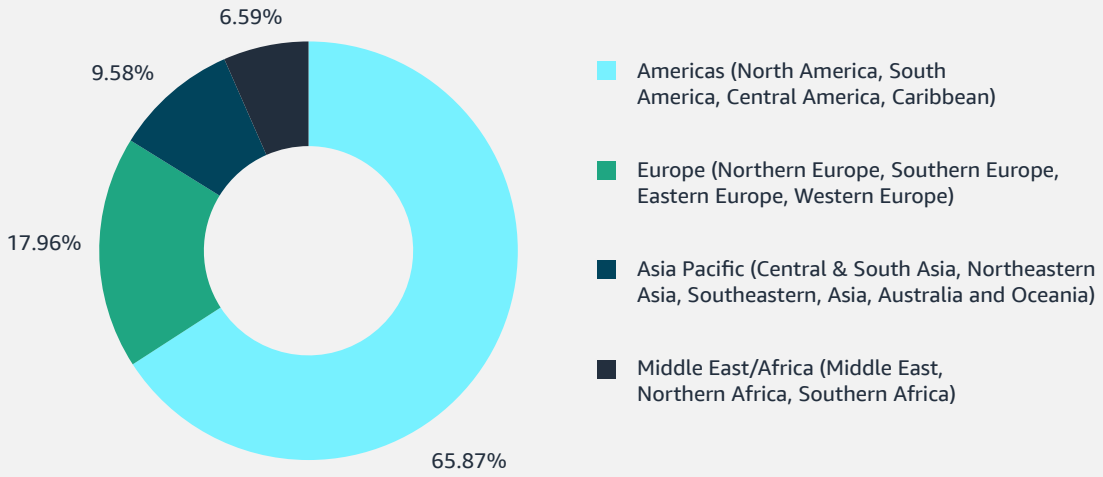
Successful CDOs see their role as making their internal clients more successful in achieving their objectives. They don't focus so much on their own performance alone but on building coalitions to help their organizations succeed.

Research respondents and methodology

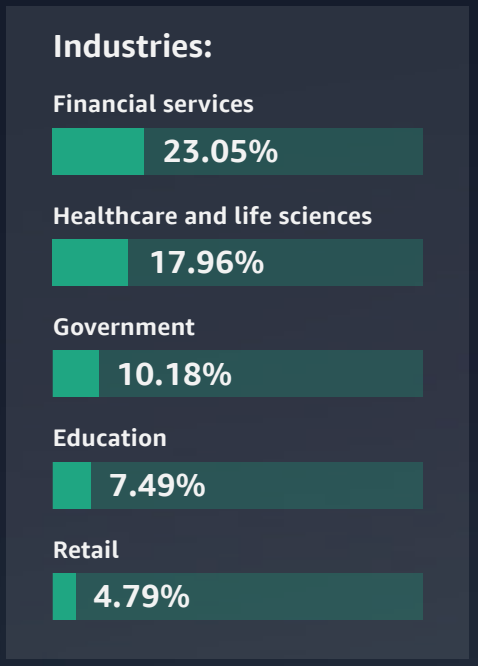
This study was conducted in the summer of 2023 and is the continuation of a similar study conducted last year. This study is sponsored by Amazon Web Services (AWS) and conducted in collaboration with the International MIT Chief Data Officer and Information Quality (MIT CDOIQ) Symposium. We conducted a global quantitative survey of 334 respondents in CDO or equivalent titles and level positions. In addition, qualitative interviews were completed with 12 CDOs. The quantitative survey included the Americas, Asia/Pacific, Europe, and Middle East/Africa regions. The largest industry represented in the survey was financial services, with 23 percent of respondents, and healthcare, with 18 percent.

The quantitative research questions addressed CDOs' demographic information, number of and tenure in CDO jobs, professional backgrounds, key activities, challenges, perceptions of how successful they are in their jobs, what they do to create value for their organizations, and how they're preparing their organization to leverage generative AI. The qualitative interviews explored similar topics but went into greater detail on organizational structures, key initiatives, generative AI and ways to create value as a CDO.

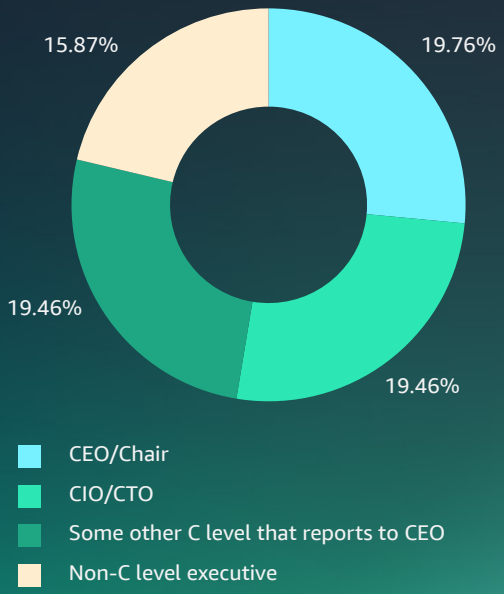
Demographic snapshot



Industries:



Who do CDOs report to?



SECTION 1

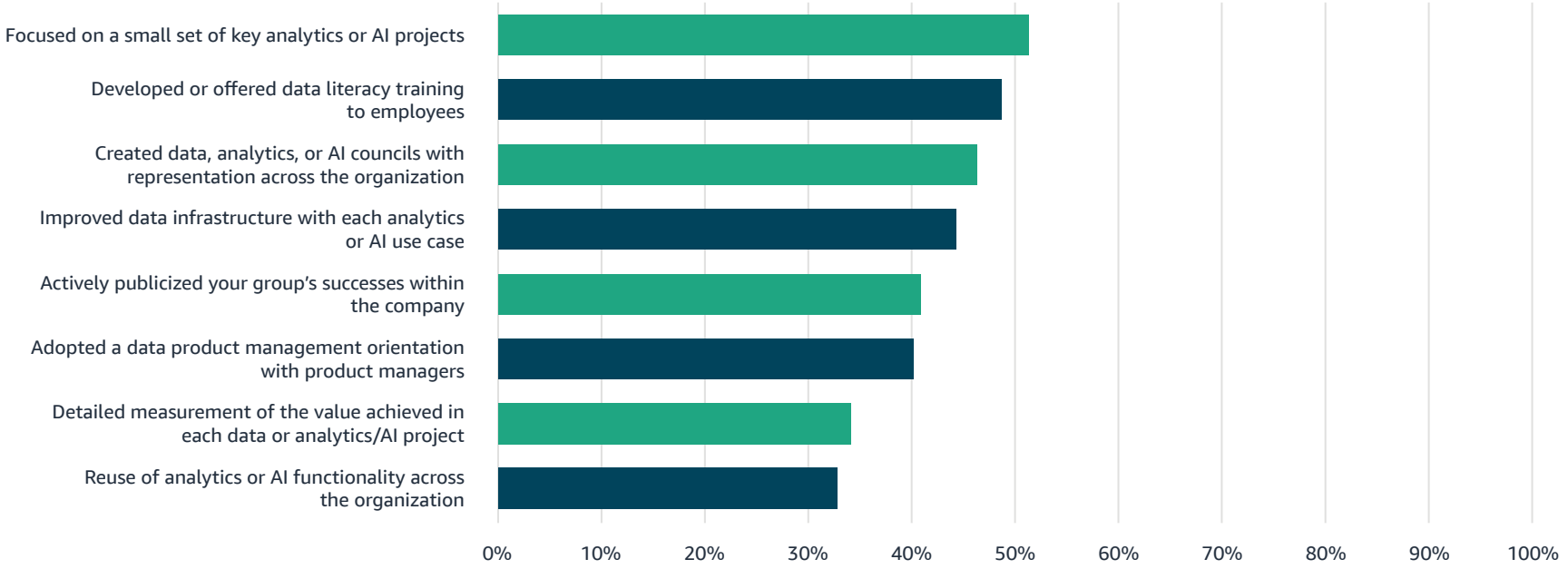
Striving for visible business value creation

CDOs are focused on creating visible value for their organizations. In the 2022 research, CDOs discussed the value of analytics and AI projects as a means to deliver solid and measurable value. In 2023, "analytics and AI" continue to reign supreme. In this year's survey, over half of the respondents indicated that they focus on delivering a "small set of analytics or AI projects" as a value creation approach. Nearly half of them have instituted data literacy training and have organized data, analytics, or AI councils.

Although data management is one of the primary responsibilities for CDOs, 44 percent say they focus on data management initiatives such as improving data infrastructure in the context of each analytics and AI use case and not as a separate initiative. Forty percent employ a "data product management" approach—treating data like a product and strategically developing, launching, supporting, and ensuring success of data products within the organization for data-driven analytics and AI products for employees or customers.

Chart 1

Which of the following initiatives have you undertaken to bring more value to your organization as CD(A)O?



CDO's elaborated on different aspects of creating value for their organizations in one-on-one interviews.

"We are use case-driven, so for us, the first two priorities were revenue growth management—pricing and promotion in store and online—and marketing and media effectiveness analytics. We're trying not to take the "boil the ocean" approach to fixing our data. If we worked on everything at once and we talk five years from now, I still would say we're still working on trying to fix our data."

Diana Schildhouse
Chief Data, Analytics, and Insights Officer, Colgate-Palmolive

"The data product focus has brought data and analytics people much closer to the rest of the organization. Now data product managers will start to follow the same way of working as the PMs building customer facing software and I have taken responsibility for technology as well as data."

Sebastian Klapdor
Chief Data and Technology, Vista

"People in this job talk about metrics and value, but I think the real value add is making other people successful. This is a client service role; you have to make the other person successful… I don't have numbers to make. It's their numbers, not my numbers."

Naras Eechambadi
Chief Global Data and Analytics Officer, Universal Music

SECTION 2

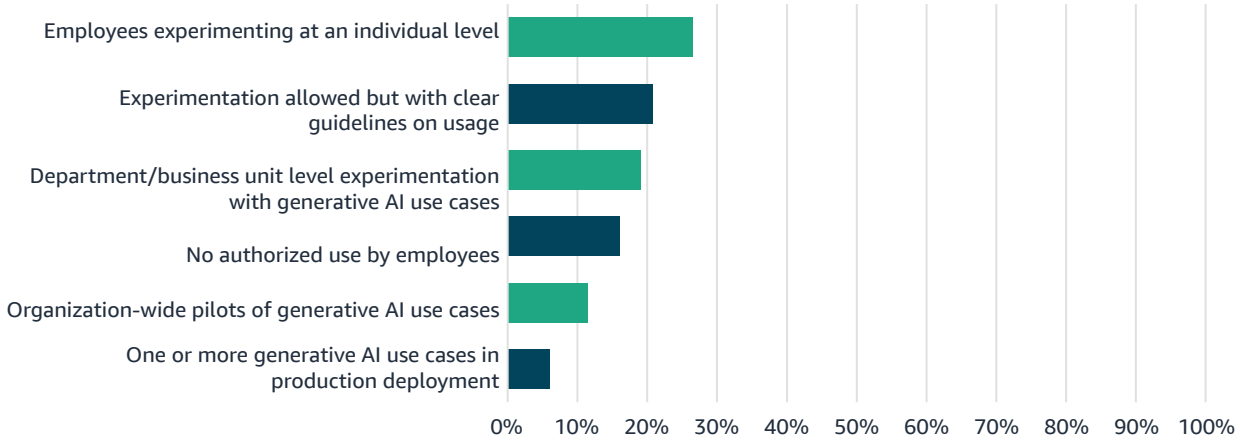
Generative AI: The latest technology to create business value

Embracing new technologies is, of course, one approach to adding value. Many CDOs reported that they are responsible for generative AI. And while it is still early days for generative AI adoption, it's clear that CDOs believe it is an important technology with an eventual profound influence on their organizations. The timeframe and paths to adoption will become clearer in the months and years ahead. At this moment, generative AI use is largely experimental. But that hasn't held back the excitement and attention. Some CDOs feel they have benefitted from the halo effect. One commented that generative AI has made her "the belle of the ball."

When asked how their organizations were addressing generative AI overall, the largest category, nearly one third, said they were "experimenting at the individual level," rather than creating enterprise strategies. An additional 21 percent indicated that they were experimenting, but with usage guidelines for employees. A slightly smaller percentage, 19 percent, had departments or business units that are experimenting with use cases. Only 6 percent had a generative AI application in production deployment, and 16 percent noted that their organizations had banned use by employees.

Chart 2

How is your company addressing generative AI?



80%

CDOs agree that "generative AI will transform my organization's business."

Despite the exploratory nature of most organizations' generative AI initiatives, CDOs have bigger plans for the future. Eighty percent agree that "generative AI will transform my organization's business (although 45 percent of that group only "agree somewhat" with the assertion). Forty-six percent "foresee or are already seeing wide adoption of generative AI." Sixty-two percent say they are "planning on investing more in generative AI."

Universal Music has recently entered into an agreement with YouTube to explore how generative AI can be used in music and music-oriented videos in ways that protect the intellectual property rights of artists.

One CDO who agrees with the transformative nature of the technology is Naras Eechambadi of Universal Music. His company's strong interest in the technology is not surprising, since generative AI can create music, write lyrics, and imitate the voices of artists.

"Our clients tend to be on the cutting edge, and we are a very innovative company. We're all very interested in generative AI. Clients and employees are experimenting within the guardrails that have been placed on them. And I expect that it will gather steam slowly and there'll eventually be some kind of tipping point. We'll suddenly realize, wow, this thing has had a tremendous impact."

Naras Eechambadi
Chief Global Data and Analytics Officer, Universal Music

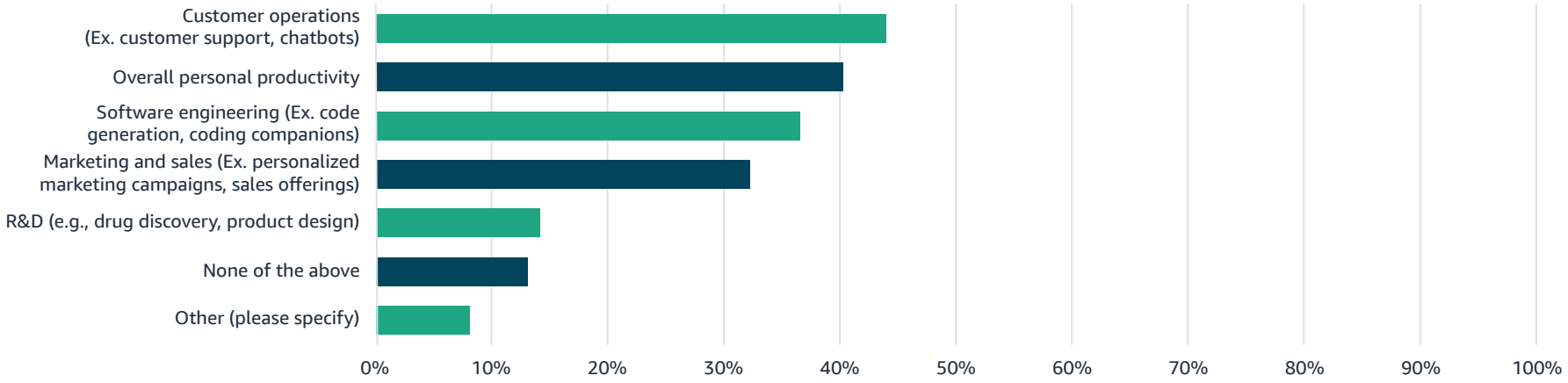
SECTION 2.1

Top use cases of generative AI

In terms of specific use cases for generative AI that their companies are prioritizing, CDOs see a broad range:

Chart 3

Which generative use cases is your company prioritizing?



"We are using generative AI for synthesizing consumer and shopper insights and other marketing use cases. We're talking with our marketing agencies about mass personalization at scale of ads or offers. And the other area is innovation. So new product development, perhaps some automated consumer insights, and then automated consumer testing with digital twins.
We want to use it to understand consumer sentiment and what they're asking for. So those are the enterprise "big bets" that we're looking at—the ones that would require senior level advocacy and corporate investment."

Diana Schildhouse

Chief Data, Analytics, and Insights Officer, Colgate-Palmolive

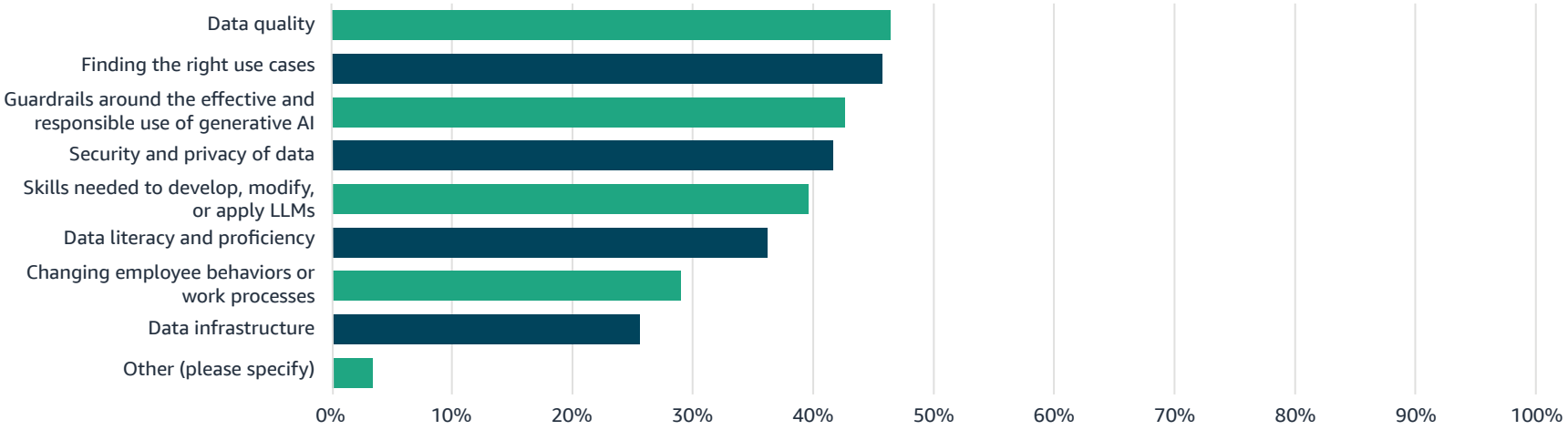
SECTION 2.2

Challenges with generative AI

Succeeding with generative AI at an enterprise level is challenging, however, when asked what they find the greatest challenges to realizing the full potential of the technology, nearly half (46 percent) of CDOs pointed to data quality and finding the right use cases. Establishing guardrails for responsible use (43 percent) and security and privacy of data (42 percent) were slightly lower concerns.

Chart 4

What is the biggest challenge for your organization in realizing the potential of generative AI?



Some companies have also made progress in establishing guardrails for responsible use. JoAnn Stonier, until recently Mastercard's chief data officer and now Mastercard fellow of data and AI, said in an interview that generative AI requires a different approach to governance:

"For traditional AI, we built a robust methodology and governance process involving understanding the data, understanding the models, and reviewing their outcomes. For generative AI, we have to focus heavily on how the query is defined, on prompt engineering, and the use of proprietary data in prompts. Both the models and the outcomes are more difficult to understand. Guidelines for generating code need to be different from those for generating marketing copy. We've engaged lawyers, HR professionals, engineers, and system and data architects in creating our approach to generative AI."

JoAnn Stonier
Fellow of Data and AI, Mastercard

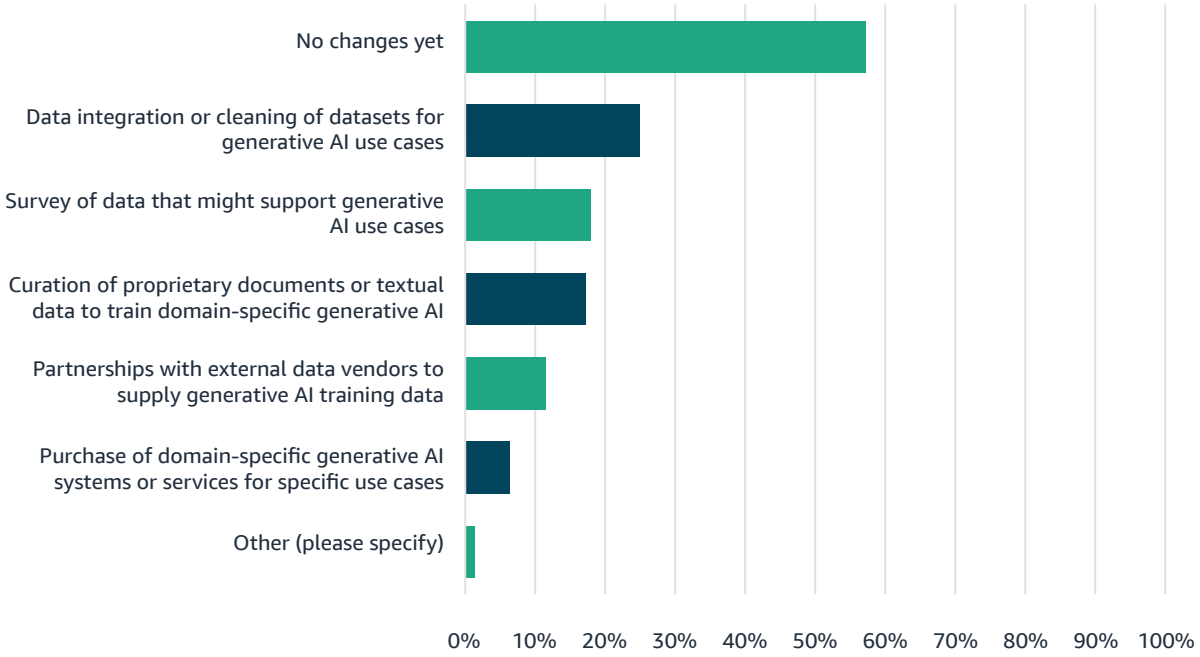
SECTION 2.3

Data strategy for generative AI

Although generative AI is in the early stage within large enterprises, it's becoming clear that data strategies will need to change in order to make the technology successful. When CDOs in the survey were asked how their company's data strategy has changed with generative AI, a majority, 57 percent, said there were no changes yet. But a quarter of CDOs said they were pursuing data integration or cleaning, 18 percent were surveying data to understand what might support generative use cases, and 17 percent were curating documents or text for domain-specific applications.

Chart 5

How has your data environment changed to support or enable generative AI?



93%

of CDOs agreed that data strategy is crucial for getting value out of generative AI.

57%

percent have not made any changes to their data strategy yet.

Morgan Stanley Wealth Management is one organization that has already addressed data issues. Morgan Stanley's Wealth Management practice initially identified fifty possible use cases of generative AI and has moved ahead on several. Jeff McMillan, the unit's chief data, analytics, and innovation officer, described how the company dealt with data issues for one important use case involving a knowledge management application:

"We have been curating our document-based knowledge for a while. Every single piece of research content has to be reviewed by a registered compliance person, so we know the training content is of very high quality. Even in non-research content, we have a team that scores individual submissions on issues like tagging requirements, broken links, presence of a summary up front, and we give each document a grade. We also had to spend a lot of time thinking about different content sets and optimizing the results. For me, the most important thing about these models is that they have to be transparent. The user should know this is what I put in, this is what came out, these are the documents the output came from, and this is confidence level for that question."

Jeff McMillan
Analytics and Data Officer, Morgan Stanley

McMillan also makes one other data issue clear: "These large language models do not solve the problem of disparate data sources. Companies need to address data integration and mastering before attempting to access data with generative AI."

CDOs agree that data strategy is the foundation of Generative AI. However, only 37 percent agree that "my team and I have the right data foundation to pivot to generative AI."

"Generative AI holds immense potential to transform the pharmaceutical industry, especially in R&D. There are plenty of use cases down the value chain, like enhancing our understanding of disease and drug targets, drug design, streamlining clinical study design, optimizing trial operations through the implementation of simulation and digital twins, and simplifying complex data documentation tasks such as regulatory submissions. However, to succeed, gen AI hinges on high-quality, "business-ready" data, which is guided by a robust data foundation, data governance, and standards. At Sanofi, we are actively building these core capabilities across R&D. Our successes are due to the unwavering support of our CEO, Paul Hudson, who recognized the transformative potential of AI and data early on."

Raj Nimmagadda
CDO for R&D, Sanofi

"If we want to do AI, we need to build it on concrete, not quicksand. We are getting the process and data supply in good shape. We're working on a data inventory and catalog, a data fabric with a new data structure and metadata layer, data pipelines, and ad-hoc self-service insights generation. We believe that generative AI will be a key way to create insights from data in the future."

Walid Mehanna
Chief Data & AI Officer, Merck Group

Although not every company is taking such active steps to prepare their data for generative AI, it appears that many will be doing so in the near future. Of course, generative AI is not the first priority for every company.

Are CDOs eager to jump all in into generative AI?

71 percent of the CDOs agreed that "generative AI is interesting, but we are more focused on other data initiatives to deliver more tangible value."

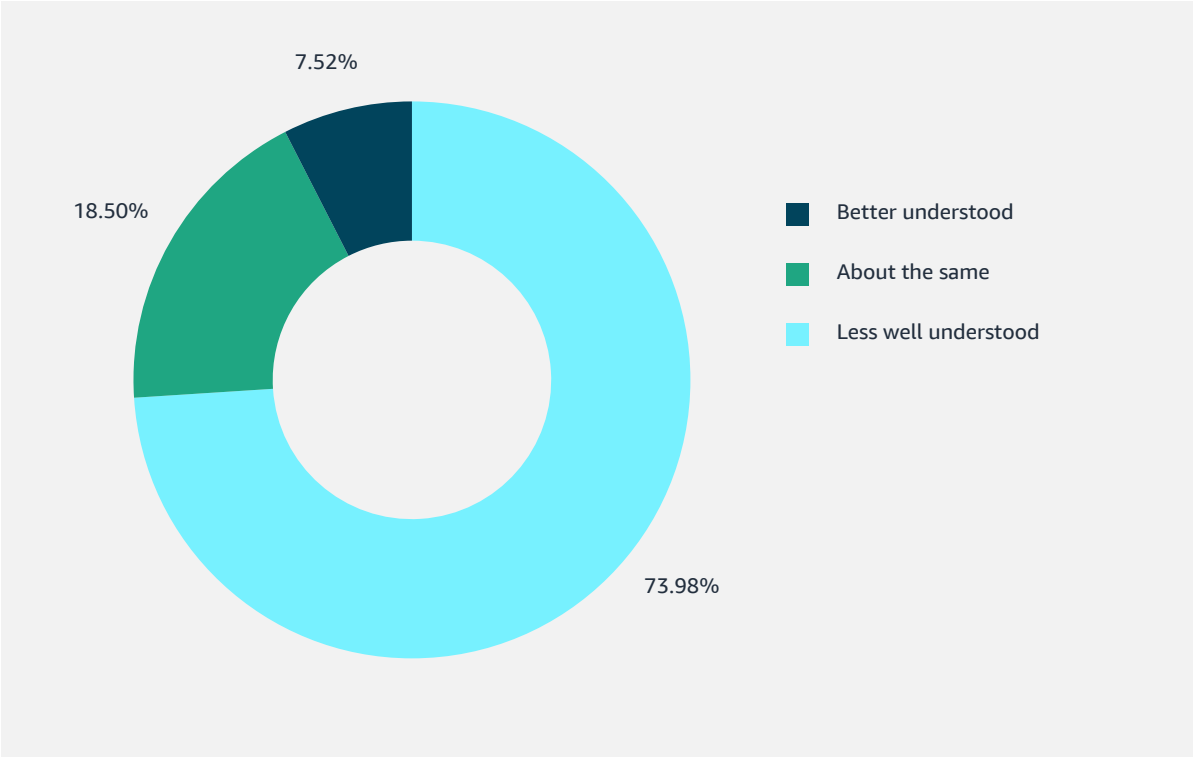
SECTION 3

CDOs' jobs are still complex and challenging

The CDO role is still relatively new among C-level executives and multifaceted. CDOs have faced high turnover in recent years and anecdotal evidence suggests that it was particularly high in 2023. While increasing numbers of companies are appointing CDOs, the job is still poorly understood for a variety of reasons.

Chart 6

How well do you feel that the roles and responsibilities of the CD(A)O role are understood compared to other C-level roles within your organization?



Three quarters of respondents (up from 62 percent last year) said that their role is less well understood than other C-level roles in their organizations.

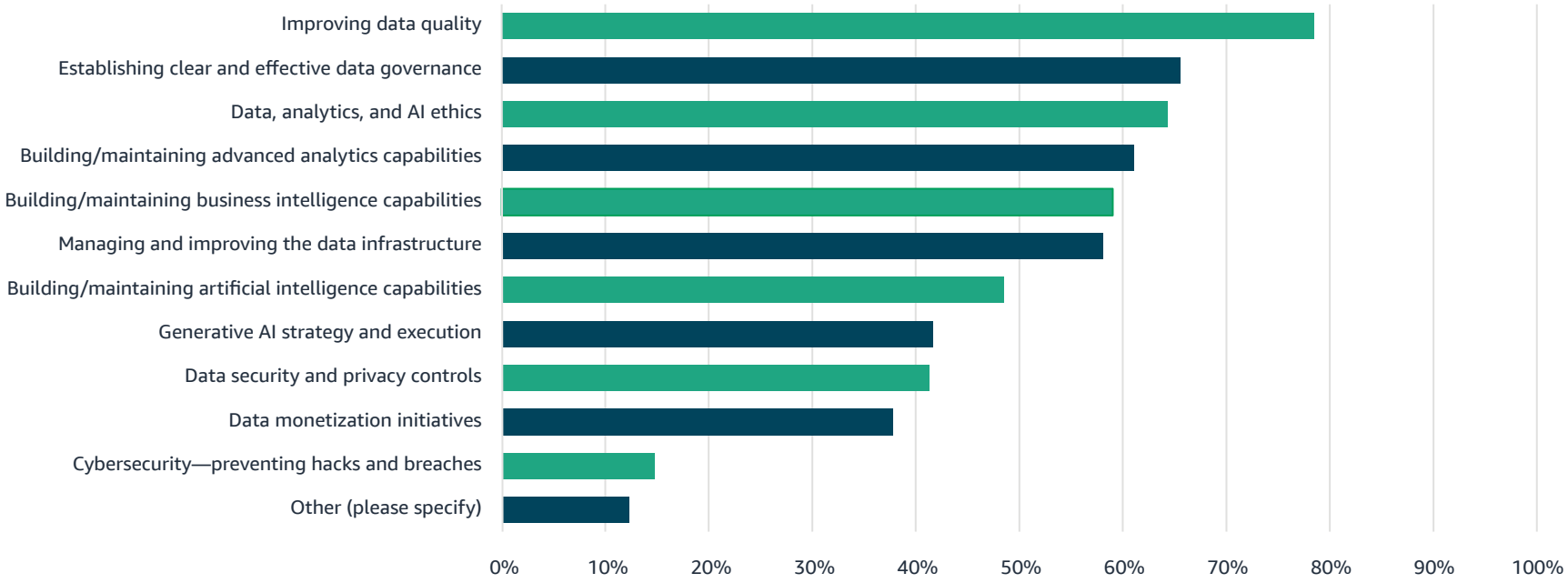
Another contributing factor to this finding could be the lack of a standard list of responsibilities. CDOs indicate a stronger focus is on business objectives. Seventy-four percent of CDOs have either a business focus or an equal business/technology focus, whereas only three percent define success in terms of technical accomplishments alone. One CDO commented in an interview:

> **"My job is not well understood; I wrestle with that issue every day. It's confusing to people since there are several groups in the organization dealing with technology and data. I have a tough audience internally, but they do appreciate some things, and I'm making an impact toward helping them understand that data is important."**

CDOs are responsible for overall data strategy, but there are many more specific tasks for which they are also accountable. The highest percentage—and up 30 percent from last year's survey—is "improving data quality," followed by "establishing clear and effective data governance," which was ranked first in the list this year at 78 percent, up 15 percent from last year. New this year, of course, is "generative AI strategy and execution." Forty-two percent of CDOs in the survey say they are responsible for it. The "other" responses included some variation of content and communication, cost management and analytics, and procurement.

Chart 7

Which of the following responsibilities are included in your job?

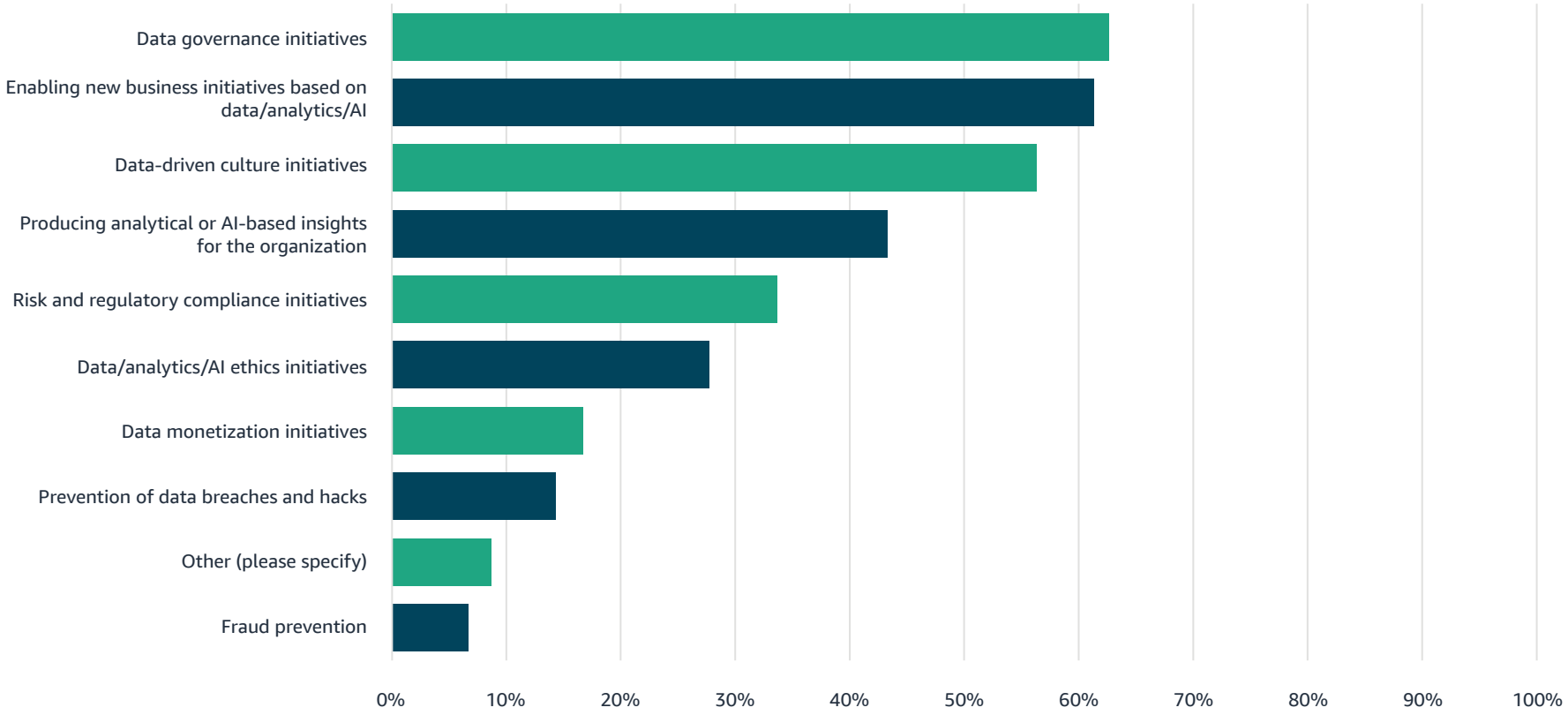


We also asked CDOs to note the activities on which they spend at least 20 percent of their time. In this list, data governance came out on top, at 63 percent. Also high on the list was "enabling new business initiatives based on data/analytics/AI," with 61 percent confirming that they spend 20 percent or more of their time on it. Producing analytical or AI-based insights for the organization was also chosen by 43 percent.

In other words, many CDOs are also effectively or officially CDAOs or CDAIOs.

Chart 8

To which of the following activities have you devoted 20 percent or more of your attention?



SECTION 4

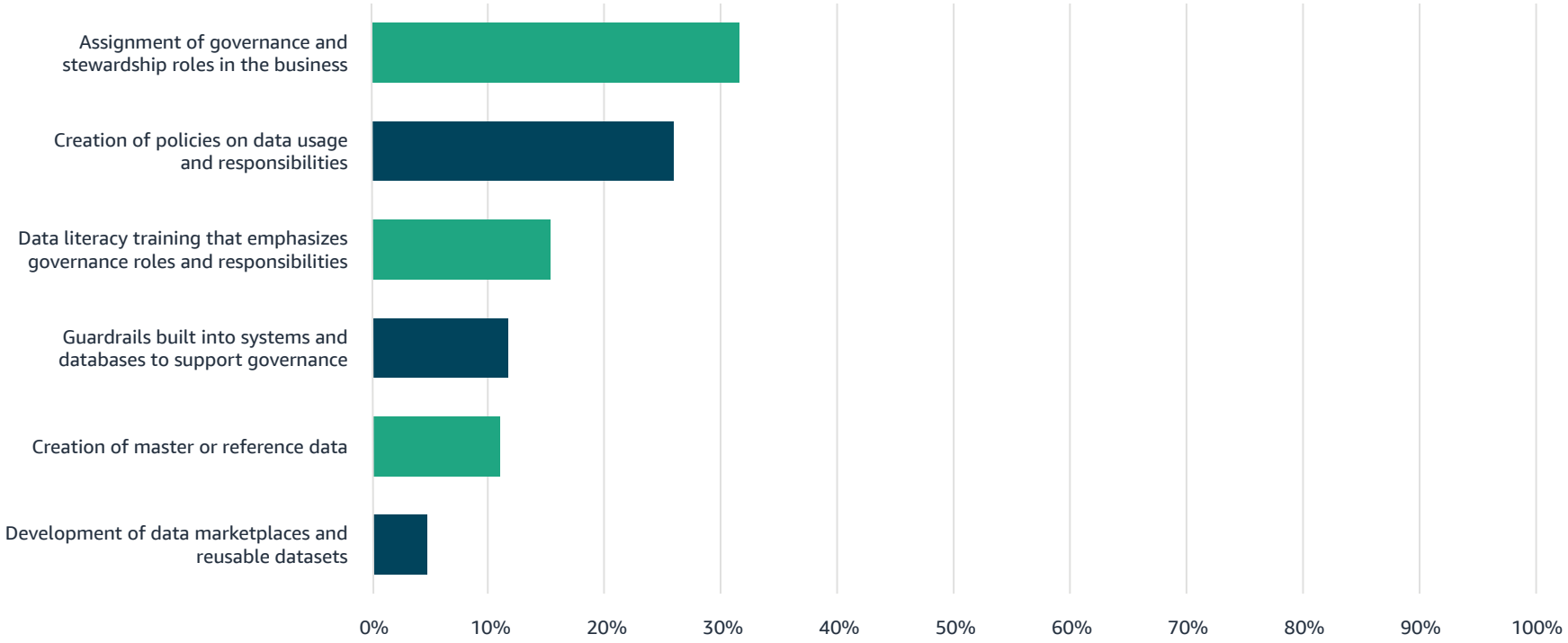
Data governance is still front and center for CDOs

Two years in a row, data governance was listed as the top activity on which CDOs spend substantial time—63 percent in 2023, up from 44 percent in 2022. CDOs stated that data governance is difficult to accomplish, in part because it involves substantial behavior change. The idea of "governance" is being supplanted in some firms by a more positive "data enablement" focus. CDOs reported a variety of approaches to governance in the survey, led by "assignment of governance and stewardship roles in the business."

"Data-driven culture initiatives" garnered 56 percent of the key activities list. As in last year's report, data monetization garnered a low level of CDO focus, at 17 percent, and fraud prevention, at only 7 percent, is probably the function of other executives.

Chart 9

What is the primary way your organization approaches data governance?



Several CDOs also mentioned in interviews that they were working on a common data platform approach for data governance to ensure that employees have access to common data.

At the Coast Guard, for example, Chief Data and AI Officer Capt. Brian Erickson is leading the creation such a platform that is modeled after ADVANA, a Department of Defense (the Coast Guard is a Homeland Security agency) centralized data and analytics platform that provides DoD users with common financial and operational data, decision support analytics, and data management tools.

At Siemens Energy, CDO Micheline Casey is creating the enterprise data infrastructure, foundations, and process to support the company, which recently spun out of Siemens and is now independent. The data platforms and processes are built to support curated data discovery with data catalogs, graph technology, and metadata, so that data can be used in a mesh construct to build data products, AI-powered digital solutions, or predictive insights. Her group is shoring up master data quality by prioritizing key data domains such as material master, business partner master, and install base master, developing standards and processes, and building data pipelines to speed access and digital development.

SECTION 5

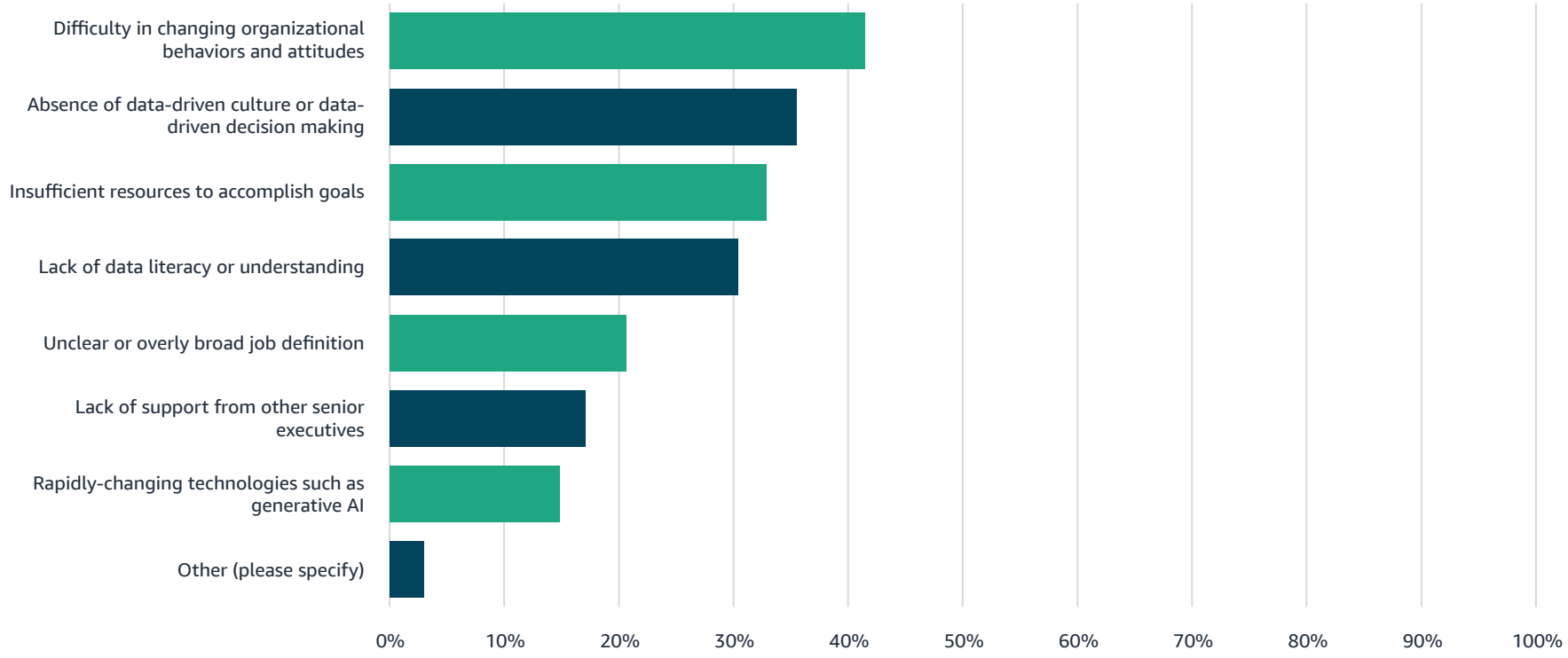
Greatest challenge for CDOs: creating a data-driven culture

Consistent with other surveys, the challenges faced by the CDOs represented in this survey were mostly organizational and behavioral, rather than technological. “Difficulty in changing organizational behaviors and attitudes” was the most frequently mentioned challenge in the role, with 70 percent facing it. The related issue of “absence of data-driven culture or data-driven decision-making” was selected as a challenge by over half of them (59 percent). Another culture-related issue, “lack of data literacy or understanding,” was cited by half of the respondents. Only “insufficient resources to accomplish goals,” at 55 percent, was near these cultural and organizational challenges in importance.

Interestingly enough, the rapidly changing technologies such as generative AI were at the bottom of the list of challenges (24 percent), which indicates that CDOs have more pressing challenges to tackle first.

Chart 10

What have been your greatest challenges in the CD(A)O role?



We saw before that more than half of CDOs devoted one-fifth of their time or more to data-driven culture initiatives. These most commonly include data literacy programs but may also encompass change management approaches for specific data or analytics initiatives. Preparing an organization for the workforce implications of generative AI was another cultural change activity mentioned in a CDO interview.

Given the organizational and behavioral challenges of making companies more data-driven, many CDOs find it important to attempt it at a pace with which their organization is comfortable.

"We have a focus on data governance and data excellence, but we target it to areas where it is needed. We also try to tie our data and analytics work to specific decisions. We call it decision intelligence. We are just trying to make it easier for the organization to do data-oriented work."

Micheline Casey
Chief Data and Analytics Officer, Siemens Energy

Will we see more CDOs focus on their data strategy or building a data-driven foundation?

It is important for CDOs to be aware of their organization's readiness for data initiatives to adopt the easiest path to achieve desired behaviors and business outcomes. Culture change is difficult both to accomplish and to measure, but the survey and interviews leave little doubt that it is necessary for CDOs and their organizations to flourish.

In conclusion

The chief data officer role—increasingly combined with analytics and AI—is one of the most rapidly-changing jobs in business. As companies move toward digital transformation and data-driven decisions, it's also one of the most central roles in that transformation. It is apparent from both the large survey of CDOs and our interviews that adding generative AI to the mix of an organization's capabilities will be one of the most pressing tasks of the next several years.

A well-crafted data strategy—often listed as CDOs' primary responsibility—is the foundation of generative AI success. CDOs will need to encourage experimentation, start with the right use cases, and treat organizational data responsibly to showcase visible business value with generative AI.

Ten keys to succeeding as a CDO

1. Constantly look for ways to add visible value to your organization.
2. Add analytics and AI to the CDO portfolio whenever possible.
3. Try to build coalitions and make other people successful in achieving their objectives.
4. Encourage experimentation with generative AI, but try also to find strategic use cases for the technology.
5. Don't abandon existing data, analytics, and AI initiatives in favor of generative AI, but add it to the mix.
6. Begin transforming and curating data, both structured and unstructured, to make it easier to succeed with generative AI.
7. Adopt a common platform for data, analytics, and machine learning features for the organization to employ in its decision-making.
8. Employ an "enablement" approach to achieving the data-related behaviors you desire, not a "governance" one.
9. Take a use case by use case approach to improving data management.
10. Strive to create a data-driven culture, but don't force changes, and take them slowly.

About the authors:

Tom Davenport is a Distinguished Professor of Information Technology and Management at Babson College, Fellow of the MIT Initiative on the Digital Economy, and Senior Advisor to Deloitte's AI practice. He's published 23 books and over 300 articles for Harvard Business Review and many other publications. He has been named among the world's Top 25 Consultants, the top three business/technology analysts, the 100 most influential people in the IT industry, and the top fifty business school professors in the world. He's worked with many of the world's leading companies on data, analytics, and AI strategies.

Richard Wang is the Founder and Executive Director of the Chief Data Officer and Information Quality Symposium and the CDO Certification Program. He has served as the Chief Data Quality Officer and Deputy CDO of the US Army, Pentagon, and the first CDO at the State of Arkansas. Dr. Wang is an internationally renowned pioneer in the field of Information Quality. He is a professor at the University of Arkansas at Little Rock; before that he served as a professor at MIT for almost a decade.

Randy Bean has been an advisor to Fortune 1000 organizations on data leadership for over three decades. He is the author of the bestselling Fail Fast, Learn Faster: Lessons in Data-Driven Leadership in an Age of Disruption, Big Data, and AI, and a contributor to Forbes, Harvard Business Review, and MIT Sloan Management Review. Randy was founder and CEO of NewVantage Partners (NVP), a data strategy advisory firm that he ran for 20 years prior to its acquisition by Paris-based global consultancy, Wavestone, in December 2021. He currently serves as Innovation Fellow, Data Strategy at Wavestone.

About AWS

For over 15 years, Amazon Web Services (AWS) has been the world's most comprehensive and broadly adopted cloud offering. AWS has been continually expanding its services to support virtually any cloud workload, and it now has more than 200 fully featured services for compute, storage, databases, networking, analytics, machine learning (ML) and artificial intelligence (AI), Internet of Things (IoT), mobile, security, hybrid, virtual and augmented reality (VR and AR), media, and application development, deployment, and management from 87 Availability Zones within 27 geographic regions, with announced plans for 21 more Availability Zones and seven more AWS Regions in Australia, Canada, India, Israel, New Zealand, Spain, and Switzerland. Millions of customers—including the fastest-growing startups, largest enterprises, and leading government agencies—trust AWS to power their infrastructure, become more agile, and lower costs.

Data is at the center of every application, process, and business decision. It is the cornerstone of almost every organization's digital transformation. But building a strategy that unlocks the value of data for your entire organization is not an easy and straightforward journey. Companies struggle to get a handle on where all their data sits, how to connect and act on that data effectively, and how to manage access to that data. And as data volumes grow, this only gets more difficult. The inability to use data effectively can hinder rapid decision-making and sustained innovation. Only AWS provides you with the capabilities you need for an end-to-end data strategy that will serve you now and in the future. AWS provides the most comprehensive set of data capabilities to give you optimal price performance for any data workload or use case; end-to-end data governance so you can free your teams to move faster with data; and integration capabilities so you can quickly and easily connect and act on all your data, no matter where it lives. And with built-in intelligence and automation in all our data services, AWS makes the complexities of data management easier, so you spend less time managing data and the underlying infrastructure and more time getting value from it. To learn more about how AWS helps organizations fuel innovation with data and AI, visit aws.amazon.com/data.

About CDOIQ

The Chief Data Officer & Information Quality (CDOIQ) Symposium has attracted thousands of CDOs, data leaders, and C-suite officers to share and exchange cutting-edge ideas and findings over the last seventeen years. It is the premier venue to accelerate the adoption of the CDO role globally across all industries. CDOIQ provides the feedback loop to advance data leadership processes and designs and push the leading edge of research opportunities.

Building upon this momentum, we have invited thought leaders and leading professionals in the data leadership field to launch a CDO Certification Program aiming at educating and certifying qualified data leaders. It is critical that government, industry, and academia understand the role and responsibility of CDO and how the CDO creates and demonstrates value. The certification program includes a blend of topics which are crucial to CDO practice, ranging from business to technology and encompassing both qualitative and quantitative aspects of CDO practice.