



# Data-powered enterprises: *The path* to data mastery

# Table of contents

**04**  
Executive summary

**06**  
Who should read this report and why?

**08**  
Introduction

**10**  
Defining the data-powered enterprise

**12**  
How have organizations progressed on their data mastery journeys?

**16**  
Organizations have improved decision-making and monetization, but data identification and access require more focus

**24**  
Generative AI is transforming the data landscape

# 32

What benefits do data masters receive?

# 38

How are data masters leading the data transformation?

# 50

Conclusion

# 51

Research methodology

# 53

Appendix



# Executive Summary

In the 2020 edition of our research, *The data-powered enterprise: Why organizations must strengthen their data mastery*, we found that organizations had made headway on data-powered decision-making. However, achieving data mastery was still a long road ahead for many. Today, nearly two in three executives agree that their organizations use activated data to introduce new products or services or to develop entirely new business models. This year's report, the next in the data-powered enterprises series, finds that, while many of the challenges we noted four years ago persist, they have diminished in intensity, as the challenges that were historically seen as big issues, have become secondary as new priorities like AI have come to the fore.

This year, we also explore how ready data foundations are to harness the power of generative AI. Only 40% of data executives state their organizations have developed

their non-technical foundations (culture, ethical guardrails, governance mechanisms, and legal and regulatory frameworks). Over half (56%) consider themselves mature on technical foundations (data, technology, infrastructure, and technical skills). We also investigate the rise of generative AI and its impact on data operations. We found that 60% of organizations have implemented pilots or launched early proofs of concept (PoCs) of generative AI initiatives based on their enterprise data. However, 75% of data executives cite large-scale deployment of generative AI PoCs as a major challenge. For instance, only 42% of data executives have the required data to train generative AI models. The scale of the business opportunity has increased for data-powered enterprises when compared to 2020 and so has the bar to become a data-powered enterprise as the breadth of use cases for data has risen.

# Executive Summary

We analyzed the organizations that participated in our research across data foundations and data behaviors and found that 17% qualified as data masters and reap higher benefits across data operations, generative AI, business, and financial metrics. 83% of data masters have high effectiveness in quantifying the value of data assets and monetizing them, compared with 61% of others.

**We highlight key areas that organizations should focus on in order to accelerate their journeys to becoming data-powered enterprises:**

- First, define the data strategy as a seamlessly unified effort between business and data executives
- Aim to enhance data synergy and democratization across the business with the data foundation necessary to implement and scale generative AI and other innovations powered by data
- Put in place the governance framework required for innovations powered by data such as generative AI data pilots and as the necessary elements to ensure trust in data operations
- Enable employees to be powered by data through upskilling and a collaborative culture
- Finally, equip for scaling on impact, with a focus on business priorities while leveraging a joint platform strategy

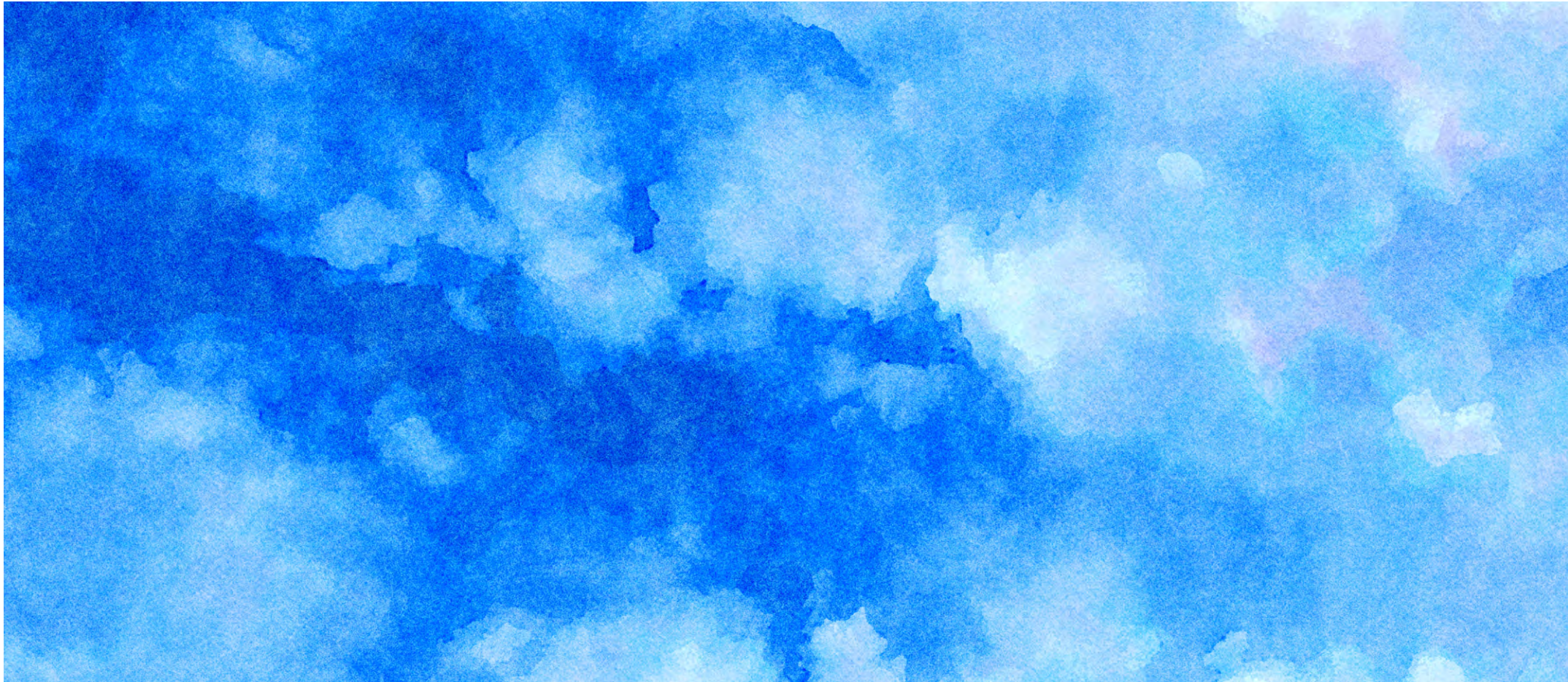


# Who should read this report and why?

This report follows the 2020 edition of our research, ***The data-powered enterprise: Why organizations must strengthen their data mastery*** and provides insights into the data-transformation journey of organizations across sectors. This report will be particularly helpful to **chief data officers, chief information officers, chief technology officers, chief analytics officers, enterprise analytics architects, and data architects**. Additionally, given the current central role of AI in the data landscape, this report will also inform AI, analytics, and data science leaders.

Finally, sustainability leaders will find it useful in light of recently introduced regulations.

This report is based on the findings of an industry survey of 500 business executives and 500 data executives from 500 organizations across 12 countries. All organizations had annual revenue above \$1 billion. Executives surveyed were director-level and above and were selected from across business and data functions. See the Research Methodology at the end of the report for more details.



# Introduction

Today data holds unparalleled potential. A small use case, such as real-time data about delayed trains, can save 27 million working hours, equivalent to €740 million in labor costs for the European Union<sup>1</sup>. In addition to improving operational efficiencies and saving costs, organizations today are increasingly monetizing data and leveraging it to boost their top lines. For example, companies like Coca-Cola are using data for intelligent revenue growth management including formulating segmentation and pricing strategies, portfolio and pack mix, and promotions<sup>2</sup>. In fact, Coca Cola's intelligent prioritization model is believed to have resulted in a revenue growth of 9% in the first half of 2023<sup>3</sup>.

However, in order to leverage data for maximum potential, organizations will need to make significant changes to

their data collection, storage, retrieval, and governance processes. Organizations that have progressed on this data mastery journey are already reaping benefits, as visible from the improvement shown in their operations and financial metrics. Further, with generative AI, the power of data has grown manifold, but do organizations have the right data foundations to be able to scale and productionize their generative AI initiatives and derive value out of those?

To find answers to such questions and to assess the current state of the organization on the data front, we conducted a global survey of 500 business executives and 500 data executives from 500 organizations across 12 countries. We also conducted detailed interviews with 10 senior executives from across industries.



# Introduction

Based on our research, we have delved into the following key areas in the report

## 01

How have organizations progressed on their data mastery journey?

## 02

Which are the key areas where organizations have improved, and have ground to cover?

## 03

Do the organizations have the necessary data foundations to leverage generative AI successfully?

## 04

What are the benefits being observed by data masters?

## 05

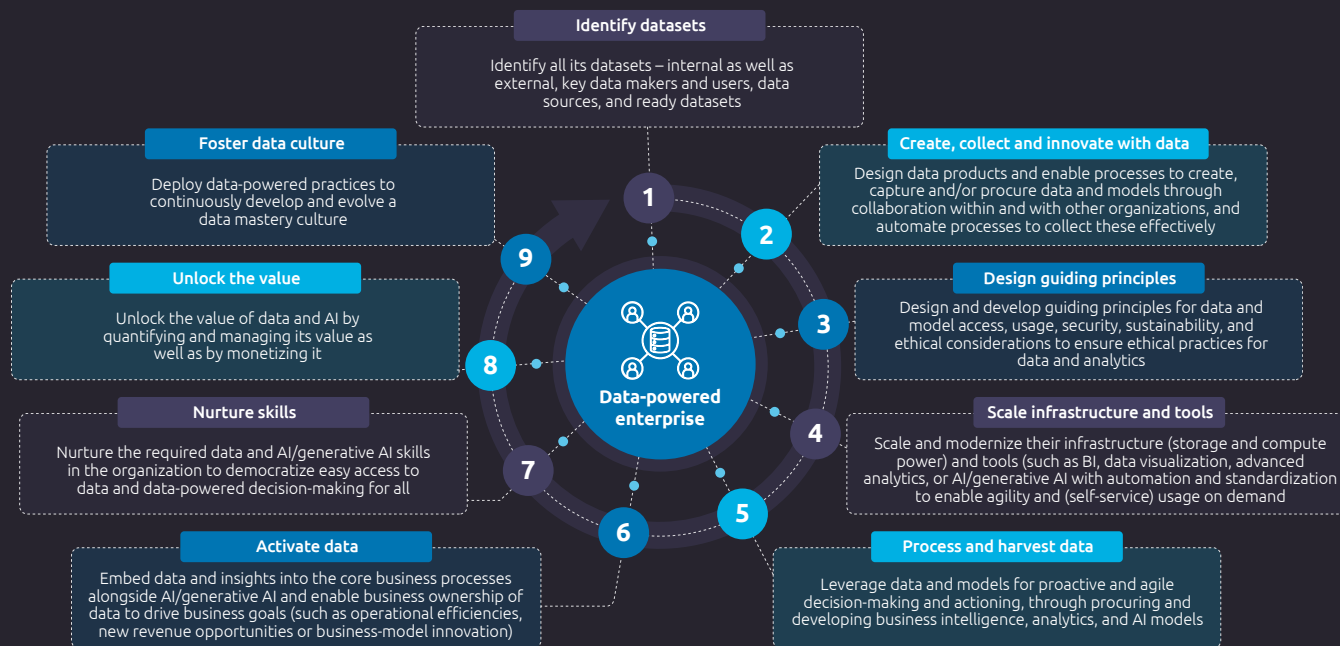
What are the key areas that organizations should focus on to accelerate their journey to becoming a data-powered enterprise?



# Defining the data-powered enterprise

We define a data-powered enterprise as an organization that can create, process, and leverage data proactively to achieve its business objectives, increase operational excellence, improve customer experience, and drive innovation.

Such an organization will be able to:



When we compare 2020 insights with 2024's, it is clear that the scale of the business opportunity has increased for data-powered enterprises, which, in turn, has meant an increase in the scale of the challenge of delivering data mastery. As we see AI adoption increase, we expect both of these bars to continue to rise, meaning that even those who are considered data-powered enterprises today cannot rest on their laurels and expect continued excellence.



# 01

**How have organizations progressed on their data mastery journeys?**

# How are organizations performing on the individual dimensions that constitute a data-powered enterprise?

We analyzed nine dimensions of data-powered enterprises (see *Defining the “data-powered enterprise”* at the beginning of the report) and found that, in the past four years, organizations have on average improved in activating data;<sup>4</sup> unlocking value from data; and scaling infrastructure, platforms, and tools (see Figure 1). Nearly two-thirds of executives state their organizations use activated data\* to introduce new products or services within existing business models or to develop entirely new business models (up from four in ten in 2020).

However, more uses for data through activation and unlocking the value have increased the foundational challenges of getting the data and identifying the right data and models. When it comes to identifying and collecting data and creating data models – also key areas for generative AI models – organizations lag. Only 42% of data executives say they currently receive the required data to train AI/generative AI models. The creating and collecting data dimension scores least well among the

nine parameters for a data-powered enterprise (see Figure 1).

The Head of AI and Automation at a telecom firm, discusses the data identification and collection challenges for large enterprises:

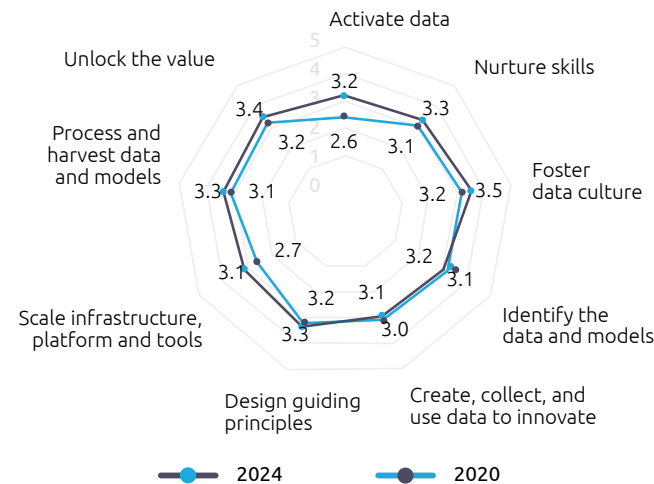
*“Many of these data-related issues are human-made, stemming from siloed operations. Individuals access data, interpret it differently, and create unique features without central coordination. This results in fragmented data processing. For example, in telcos, customer orders go through various isolated business units, including customer care, services, and fraud detection, with each team handling data separately and with minimal communication. A data-powered enterprise can only be achieved by breaking down these silos and fostering a data-driven culture.”*

Note: \*By 'activated data,' we mean embed data and insights into the core business processes alongside AI/generative AI and enable business ownership of data to drive business goals

Figure 1:

Spider chart for the dimensions of a data-powered enterprise

### Average score across nine dimensions



Source: Capgemini Research Institute, Data-powered enterprises survey, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives; Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=1,004 organizations.

## How have organizations evolved and who are the data masters?

We analyzed the organizations that participated in our research across a number of critical elements and found that 17% qualified as data masters. These elements form the two dimensions of data mastery: Data foundations and data behaviors. The data foundations

are the necessary tools and technologies with which an organization can use and leverage data, while data behaviors are part of the DNA of the organization and relate to people, processes, skills, and culture. Taken together, they drive data mastery.

### Elements of data mastery

#### Data foundation/enabler parameters

Scale infrastructure, platform, and tools

Identify all its data and models

Process and harvest data and models

Data-governance implementation

Create and collect data

Data and AI platform

#### Data behavior parameters

Activate data

Data-guiding principles (data access, interoperability, security)

Data activation vision and strategy

Foster data culture

Unlock the value

(Please refer to the appendix for additional details)

Based on these dimensions, we found:

50%

of organizations fall into the 'data laggards' category and fail to lead in either dimension (71% in the previous report)

33%

of organizations lead in one dimension but not the other (14% in the previous report)

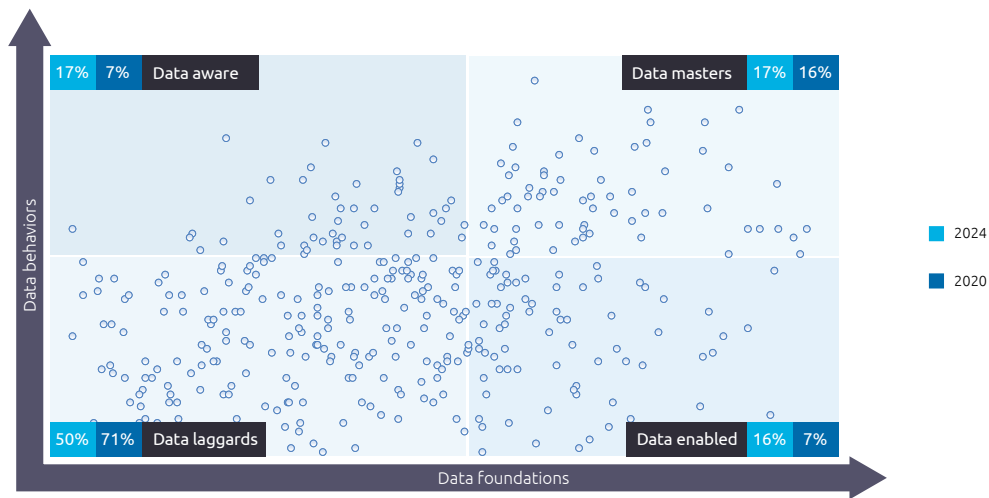
17%

of organizations fall into the data masters category, leading in both dimensions (16% in the previous report)

The share of data laggards has reduced significantly from 71% in the previous report to 50% this year, suggesting that organizations have evolved in the dimensions of data foundation and data behavior.

**Figure 2:**

One in two organizations can be classified as data laggards today



Capgemini Research Institute, Data-powered enterprises survey, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives; Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=1,004 organizations.

Note: Numbers may not add to exactly 100 due to rounding off error.



# 02

**Organizations have improved decision-making and monetization, but data identification and access require more focus**



## Since 2020, organizations have become more adept at activating and unlocking value from data for business and financial gains

Compared with a few years ago, organizations today are more effective at embedding data and insights into core business processes and using proprietary data to drive business goals. Organizations are more sophisticated in terms of unlocking and quantifying the value of data and AI and even monetizing it, either by selling to third parties or generating usable insights. The data economy in Europe was valued at around €301 billion in 2018, and is projected to reach €829 billion by 2025, from 2% to 6% of regional GDP<sup>5</sup>.

## Organizations have improved their ability to activate data for decision-making

To achieve data maturity, data must be integrated into the organization’s decision-making fabric. In our 2020 research, The data-powered enterprise<sup>6</sup>, half of the executives stated their internal decision-making was data-powered. In 2024, 60% of executives described their decision-making and actioning as data-powered (see Figure 3).

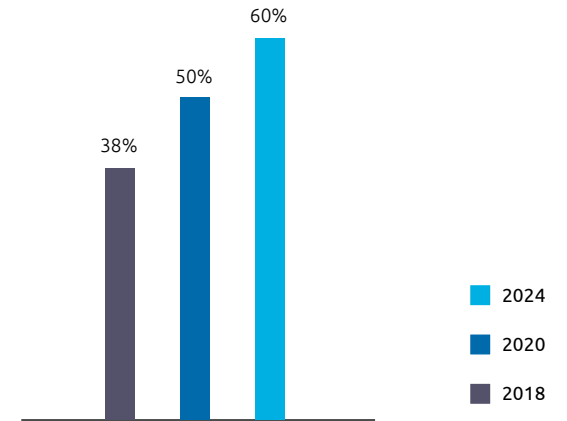
# 60%

of executives agreed that decision-making in their organizations is driven by data

**Figure 3.**

Organizations are marching ahead on data-powered decision-making

**% agreeing that decision-making in their organizations is driven by data**



Source: Capgemini Research Institute, Digital Mastery survey; April–May 2018, N = 1,338 respondents, 757 organizations; Capgemini Research Institute, Data-powered enterprises survey, August 2020, N = 1,004 organizations. Capgemini Research Institute, Data Powered Enterprise, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.<sup>7</sup>

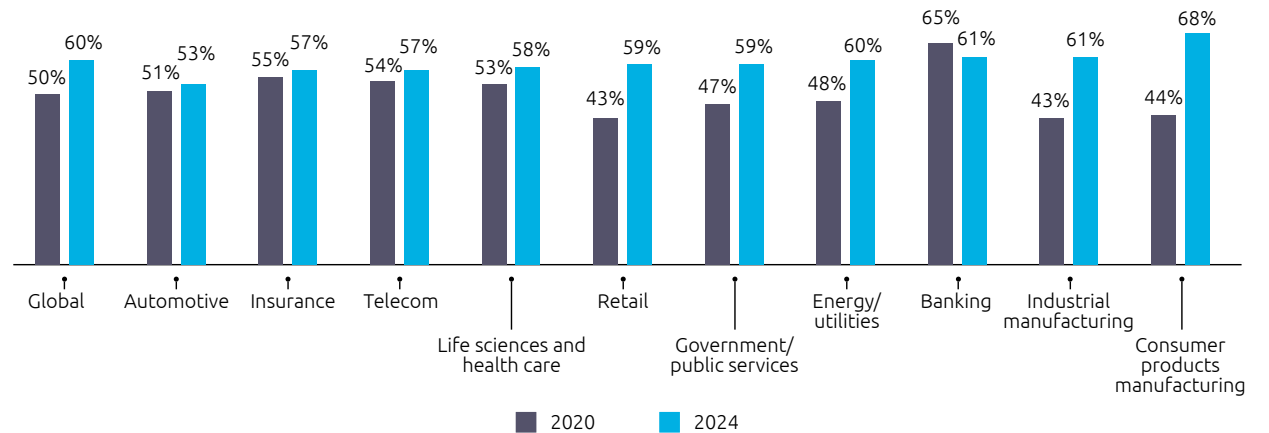
From companies like Amazon using data for one-to-one marketing, to Coca-Cola using it for enhanced customer experience, Kaiser Permanente for improving the quality of care, and Google using it to drive superior manager performance, organizations are unearthing new sources of value from data<sup>9</sup>. In fact, the consumer products industry is using data intensively, from applying consumer data to tailored marketing and designing products to consumer preferences to optimizing supply chains. The consumer products industry uses AI across several dimensions. For example, Coca-Cola installed self-service drinks machines that let consumers formulate their own drinks. Using data from this, it came up with the Cherry Sprite flavor.<sup>9</sup> Similarly, PepsiCo uses Trendscape, a tool that analyses over 500m conversations from social media platforms, news sites, blogs, forums, and reviews to address consumer preferences. During the pandemic, PepsiCo used insights from Trendscape to launch the “immunity support” version of its Propel Water.<sup>10</sup>

Sixty-five percent of business executives say their CxOs use data-powered insights to drive business. For instance, CMOs use customer data to improve customer satisfaction (CSAT) and net promoter score (NPS).

**Figure 4.**

Consumer products manufacturing lead in use of data-powered decision making

**Decision making in our organization is completely data-driven**



Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N = 1,004 organizations represented by 500 data executives and 504 business executives. Capgemini Research Institute, Data-powered enterprises survey, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.<sup>11</sup>

The CDO at a European bank, says: *“Now, leaders recognize the potential of data-driven decision-making and proactively approach data teams with specific requests, enhancing internal collaboration and aligning data strategies with organizational goals.”* Organizations are also progressing with the evolution of the data ecosystem, with 55% of data executives stating they use structured, semi-structured, and unstructured data for data-powered decision-making and implementation.

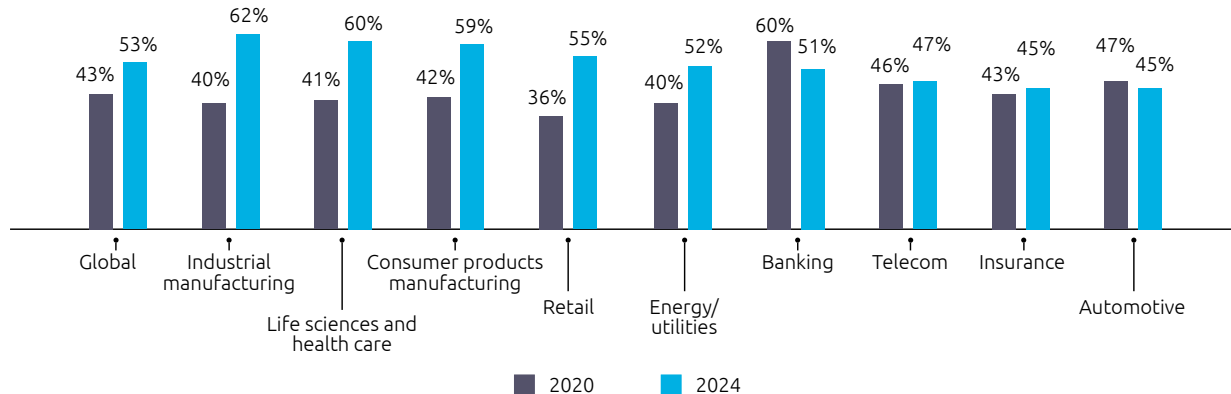
## Organizations have made major progress in unlocking value through data monetization

Nearly 70% of executives describe data as an enterprise asset compared with 62% in 2020. Further, 52% of respondents state that their organizations quantify the value of data in their accounting systems, compared with 22% in 2020. The percentage of respondents who agree that their organizations are monetizing data assets through their products and services in 2024 has also gone up (53% compared with 43% in 2020).

**Figure 5.**

Over half of executives state that their organizations are monetizing data assets through their products and services

### % of executives reporting their organization monetizes data through products and services



Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N = 1,004 organizations represented by 500 data executives and 504 business executives. Capgemini Research Institute, Data-powered enterprises survey, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.

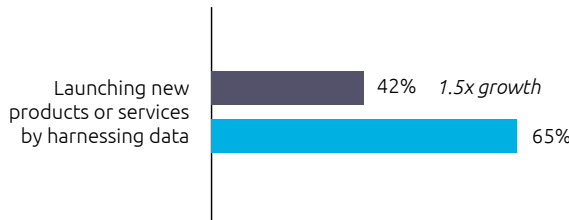
Data activation (embedding data and insights into the core business processes alongside AI/generative AI and enable business ownership of data to drive business goals) serves as a catalyst for monetization. By improving data activation, organizations can significantly boost data-revenue generation. The director of IT services at a global insurance firm, comments: *“Over the past three to five years, the volume and diversity of data generated by our organization have grown exponentially due to the proliferation of IoT [Internet of Things] devices, media, sensors, and other sources. This increase prompted us to adapt and upgrade our data models.”*

Today, nearly two in three executives agree that their organizations use activated data to introduce new products or services within existing business models or to develop entirely new business models; only about 40% were doing this in 2020. This is a massive transformation in four years and shows that data has really shifted from being a niche driver of new business to a core driver. For example, two major automotive associations, CLEPA and ACEA, have defined a “minimum viable dataset” of 42 items that original equipment manufacturers (OEMs) should cover. These range from basic information, such as odometer value and GPS position, to specifics, such as brake-fluid level warning and time to next service. This can feed into data-powered software optimization and allow the provision of advanced features like remote monitoring of vehicle health, and sending warnings and reminders to owners.<sup>12</sup>

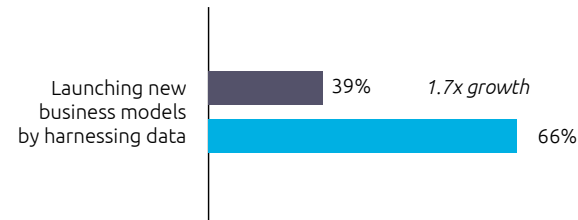
**Figure 6.**

Two in three executives agree their organizations are cashing in on data

**Percentage of executives**



**Percentage of executives**



■ 2020 ■ 2024

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N = 1,004 organizations; Data-powered enterprises survey, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.

| Industry  | Examples   |
|---|--|
| <ul style="list-style-type: none"> <li>■ <b>Retail</b></li> </ul>     | <p>Retail media networks (RMN) utilize retailers' data assets, such as first-party shopper data, to inform advertisements across a range of media. These also provide alternative revenue streams for retailers. US retail media ad spending is expected to hit \$54 billion in 2024. Walmart has launched Walmart Luminata, a new business that analyses and derives insights from customer behavior, perceptions, and channel performance, which it offers to brands, CPG suppliers, and internal merchants. Walmart's global advertising business, which includes its RMN, Walmart Connect, grew 28% to reach \$3.4 billion in 2023.<sup>13</sup></p> |
| <ul style="list-style-type: none"> <li>■ <b>Banking</b></li> </ul>    | <p>JPMorgan Chase started Chase Media Solutions, a new business unit, which directs adverts tailored from payments data to customers on its mobile app and website. It has already run pilot programs for Air Canada, Solo Stove, Blue Bottle, and Whataburger, driving incremental sales and new-customer growth.<sup>14</sup></p>  |
| <ul style="list-style-type: none"> <li>■ <b>Insurance</b></li> </ul>  | <p>The insurance industry uses telematics to personalize offerings based on usage and behavioral patterns. Progressive Insurance's Snapshot app or plug-in device collects motoring data from participating customers, allowing Progressive to offer customized discounts and incentives for safe driving behaviors. This data-powered pricing model enables Progressive to assess risk more accurately, benefiting both the company and its customers. On average, drivers who save with Snapshot save \$231 a year.<sup>15</sup></p>   |
| <ul style="list-style-type: none"> <li>■ <b>Telecom</b></li> </ul>    | <p>Telcos have made data monetization an integral part of their growth strategy, increasing their average revenue per user (ARPU) through data-powered customization and selling data to third parties. T-Mobile saw a 400% increase in customer acquisition by sending personalized videos based on these data insights. The personalized videos include the name of the customer, and have different content by target group, resulting in conversion rates that are higher than those for traditional digital marketing.<sup>16</sup> Recently, it also started its RMN leveraging 11,000 stores in the US.<sup>17</sup></p>                          |
| <ul style="list-style-type: none"> <li>■ <b>Automotive</b></li> </ul> | <p>According to our research, in 2022 there were 57 million connected cars, generating €19 million in revenue. We anticipate this number to grow to 232 million connected cars, generating €800 million by 2030.<sup>18</sup></p>  |
| <ul style="list-style-type: none"> <li>■ <b>Services</b></li> </ul>   | <p>Uber gathers data from its diverse services, including ride-hailing, courier, food delivery, and freight transport to optimize delivery routes, predict food preferences, and enhance freight logistics. By analyzing rider preferences and behaviors, Uber also provides personalized promotions and advertisements, benefiting both riders and advertisers.<sup>19</sup></p>  |

## Organizations have strengthened their foundations around infrastructure, tools, and platforms since 2020

Cloud-based infrastructure supports scalability, flexibility, and speed of decision making. The right platforms and tools for data collection and analysis facilitate democratization of findable, accessible, interoperable, and reusable (FAIR) data. Organizations have shown significant improvement on these fronts. During our last

research we had found that 51% of executives state that organizations have expanded data, BI and analytics in the cloud, while this year it is 60%. Similarly, today 76% of executives state that organizations have invested in data analytics tools as compared to 54% in 2020. There is a significant growth in investment in analytics, tools and platform, but a slowing growth in cloud adoption for data, BI and analytics.

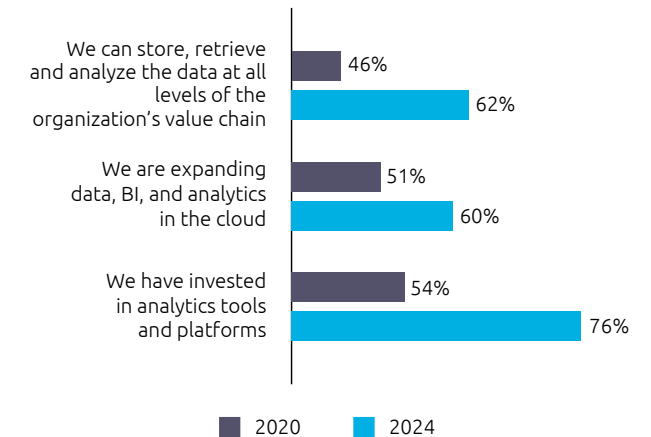
A senior executive at a European bank, adds: *“This modernization allows organizations to manage larger data volumes, harness unstructured data, and accelerate AI development. While not necessarily cheaper, the modernization of data technology allows us to extract more value and insights. The migration to cloud and adoption of modern data stacks have propelled us into a new era of data utilization.”*



**Figure 7.**

Organizations have strengthened infrastructure, tools and platform foundations

### % of executives agreeing with the following statements



Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N = 1,004 organizations; Data-powered enterprises survey, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.

## Organizations must increase focus to identify and access the right datasets

The amount of data created, captured, copied, and consumed globally in 2025 is expected to be nearly three times as much as in 2020.<sup>20</sup> Further, 80-90% will be unstructured, in the form of videos, emails, images, HTML content, and social media. From a data-management standpoint, unstructured data adds to the challenges of integration with structured data and of storage-data redundancy.

Identifying, creating, and collecting data has thus come up as an area of improvement for a large percentage of organizations today. The Chief Data Officer at a European bank, adds, *“Knitting unstructured data is going to be a big challenge in the foundational space when it comes to generative AI.”*

Further, our 2020 research revealed significant gaps in executives’ perceptions of their organizations’ data and data management. In 2024, while nearly 60% of data executives claim to have automated their data collection, only 41% of business leaders agree. Similarly, 71% of data executives are confident of the visibility of their organization’s data inventory, compared with only 45% of business executives.

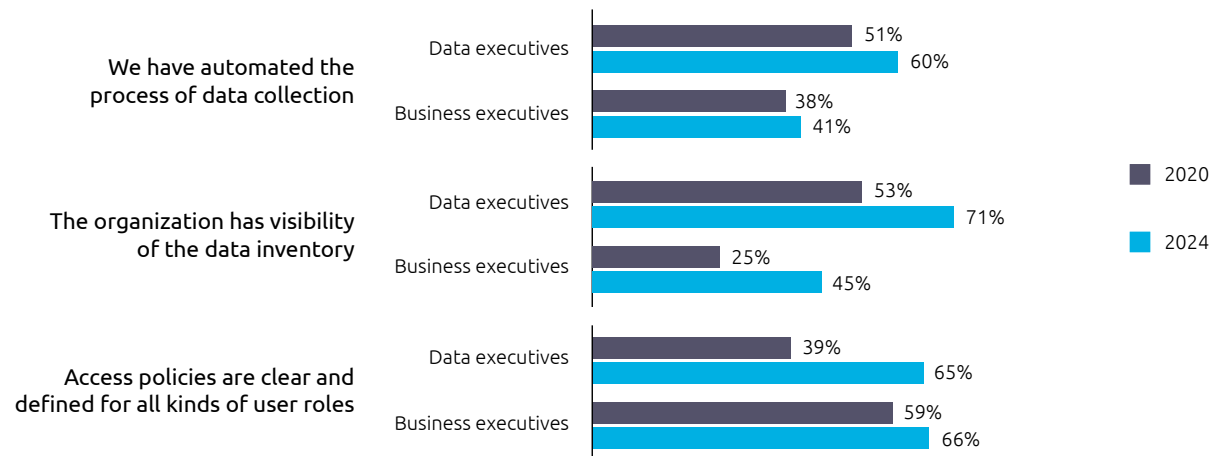
However, both data and business executives have aligned on access policies, with two-thirds of business and data

executives stating access policies are clear and defined for all kinds of user roles.

**Figure 8.**

Over 70% of data executives report visibility of their organization’s data inventory, but only 45% of business executives agree

### Share of executives who agree with the statements below



Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N = 1,004 organizations; Capgemini Research Institute, Data powered enterprises survey, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.<sup>21</sup>

# 03

## Generative AI is transforming the data landscape



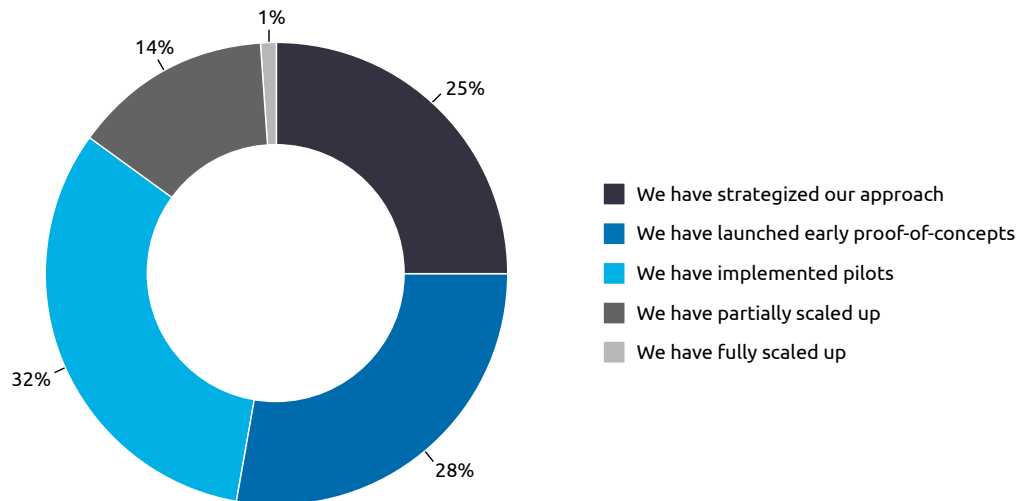
## Most organizations are exploring the use of enterprise data for AI/generative AI

Organizations are realizing the transformative power of their enterprise data through AI/generative AI. For example, Procter & Gamble rolled out chatPG, its internal generative AI tool last year to unlock the value residing in its internal data. Procter & Gamble Chief Information Officer Vittorio Cretella commented, *"We are more than experimenting. We have more than 35 use cases where that model is being complemented with internal data."*<sup>22</sup>

**Figure 9.**

Three out of five organizations have launched PoCs or pilots of generative AI initiatives using enterprise data

### What is your current state of generative AI adoption?



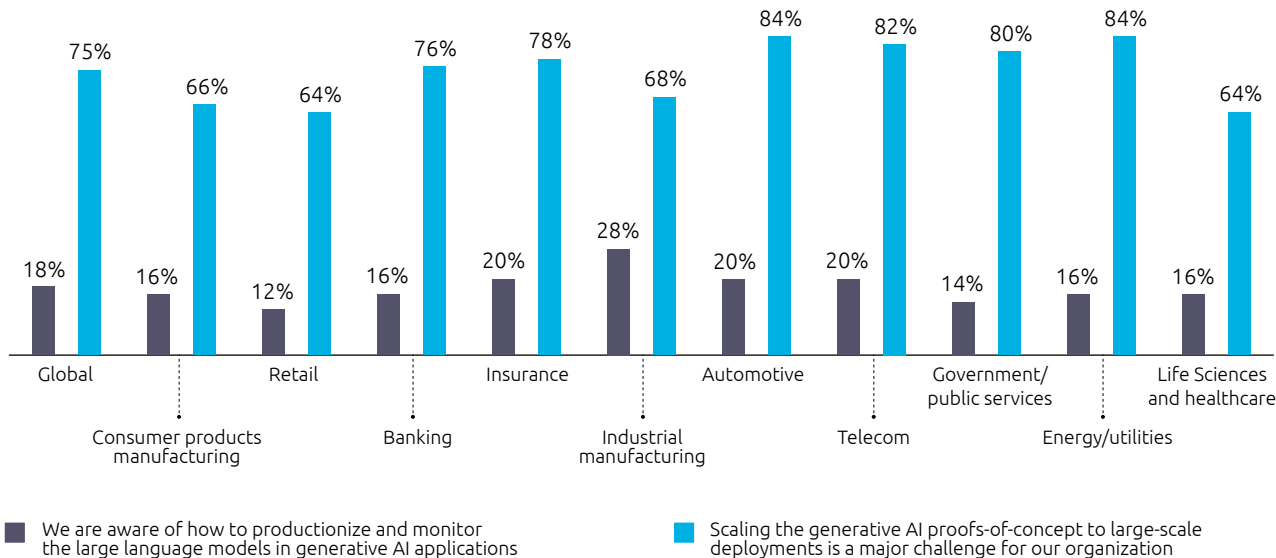
Source: Capgemini Research Institute, Data Powered Enterprise, Business Executives Survey, April 2024, N = 500 organizations represented by 500 business executives.

Around **60%**

of organizations have implemented pilots or launched early PoCs of generative AI initiatives using enterprise data (see Figure 9).

However, organizations are in the very early stages of adoption, and three-quarters of data executives agree that scaling generative AI PoCs to large-scale deployment is a major challenge.

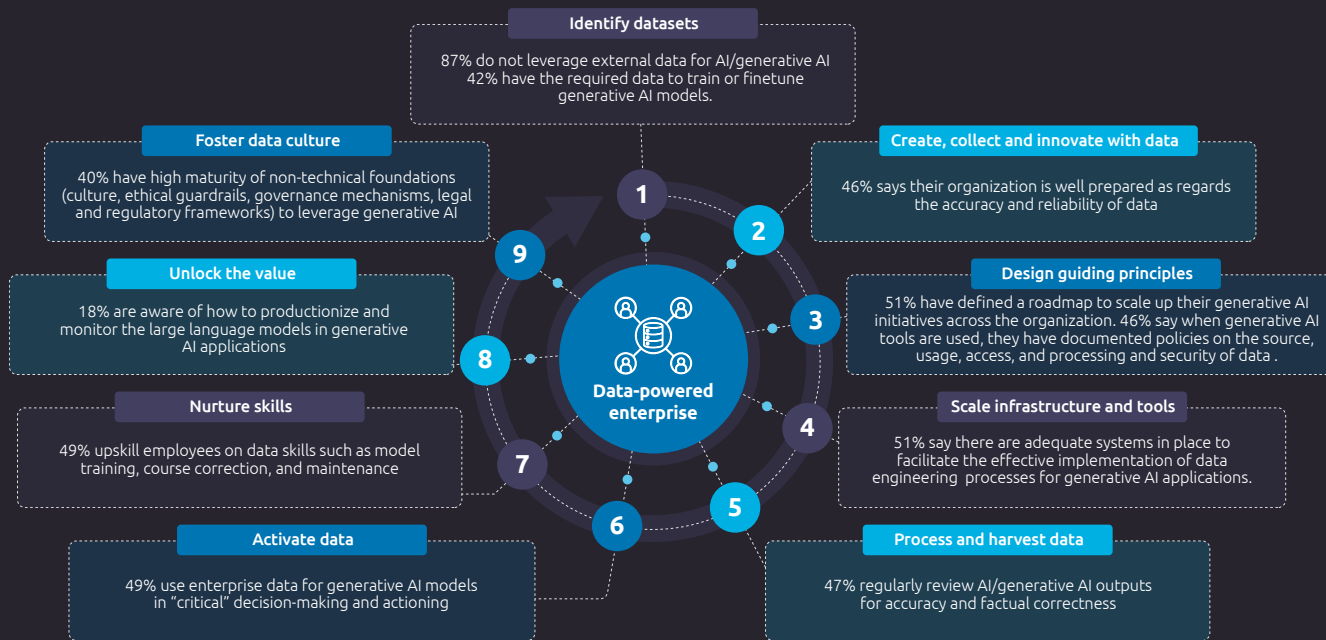
**Figure 10.** Organizations are struggling to scale up their generative AI initiatives



Source: Capgemini Research Institute, Data-powered enterprises survey, Data Executives Survey, April 2024, N = 500 organizations represented by 500 data executives.

# To harness AI/generative AI, data foundations require greater focus

Generative AI has the potential to derive value from data in an unprecedented way. However, currently, only 40% of data executives believe that their organizations are sufficiently mature on non-technical foundations such as culture, ethical guardrails, governance mechanisms, and legal and regulatory frameworks to harness generative AI. Over half (59%) consider their organizations mature on technical foundations such as data, technology, infrastructure, and technical skills. Nearly three out of five data executives agree that integrating generative AI will require significant changes to their data collection, storage, retrieval, and governance.



**Figure 11.**

There is a lot of ground to cover across various dimensions of data maturity before organizations can embrace AI/generative AI

Source: Capgemini Research Institute, Data-powered enterprises survey , Data Executives Survey, April 2024, N = 500 organizations represented by 500 data executives.

## Identifying data and models is crucial for generative AI use

There is no AI without data. However, organizations have multiple avenues to improve data and model identification.

87%

of data executives agree that their organizations do not use external data for AI/generative AI.

42%

of data executives have the required data to train generative AI models.

49%

of data sources are in the cloud, while the remainder still reside in local servers, leading to access challenges.

51%

of data executives say that their organization has clear processes to manage siloed data and data integration across functions.

55%

of data executives are tapping into structured, semi-structured, and unstructured data for data-powered decision-making and implementation.

## Few organizations nurture data skills for generative AI

Adopting generative AI and deriving value from it at enterprise-scale requires specialized technical and non-technical talent. But, 61% of data executives say their organization faces a shortage of skills and are looking at contract workers to fill the talent gap. Upwork, the freelancing marketplace, recognized AI as the fastest-growing category for 2023.<sup>23</sup>

With generative AI, retaining talent could be a challenge, especially if organizations fail to provide the necessary

support in terms of tools and training. Organizations still lag in these aspects. Only half are upskilling their employees on critical generative AI skills.

## A sound process for obtaining trustworthy data is vital

Ensuring the reliability and correctness of input and output data is critical to the success of generative AI. However, only 49% of data executives agree that their organizations are well prepared in this respect. There is a perception gap regarding transparency of data, with 87% of business executives believing their organizations are prepared on this front, whereas only 37% of data executives agree. As many as 61% of business executives cite the accuracy and impartiality of input and output data as major concerns.

Organizations must also address accountability of output data. Only 47% of data executives regularly review AI/generative AI outputs for accuracy and factual correctness.

Only 41% of business executives agree that their organization uses tools that help in the audit of the algorithms, and only 44% state that their organization has processes in place to address algorithmic bias.

It is crucial to understand the potential risks associated with generative AI usage. The Head of Data, Analytics and AI at a global apparel firm, adds: *"We have launched a generative AI tool for our store associates, who can ask specific questions in response to customer queries. What should I do if a customer wants to return something without a receipt? Generative AI provides the answers, with links to the internal documents that shaped that answer."*

## Enterprise AI adoption requires trust and guiding principles for data management

Organizations lack clear generative AI strategies, especially with regard to scaling up and addressing legal and regulatory aspects. Only 51% data executives say that they have defined a roadmap to scale up their generative AI initiatives across the organization. Further, only 54% of data executives are aware of the data foundations required to win in the AI era.

On a behavioral dimension, organizations must ensure, that they build trust from their employees towards their AI solutions. 56% of business executives rate their

enterprise as ready for at-scale AI adoption, while 67% of data executives agree.

Organizations must ramp up data governance, too. 56% of data executives state that their organizations are prepared to deal with potential legal complications arising from data use in AI/generative AI applications, and only 46% have documented policies around the sourcing, usage, access, processing, and security of data in the context of generative AI.

### Guiding principles around cybersecurity, data privacy, and protection topics are essential, as:

- 48% of business executives rate data protection and privacy concerns as among the biggest risks associated with generative AI. Failing to adequately isolate large language model (LLMs) from the outside world can expose internal data, leading to unauthorized access or accidental leaks.
- 46% of business executives rank cybersecurity as one of the top challenges for generative AI implementation. Malicious actors can exploit LLM capabilities through prompts to retrieve this information and inadequately secured AI systems are a soft target for cyberattacks.

- 40% of business executives identify IP infringements and copyright issues as top data risks associated with generative AI. If the AI models are trained on copyrighted or proprietary code or data, it can lead to legal complications.



# Where are organizations implementing generative AI in their data initiatives?

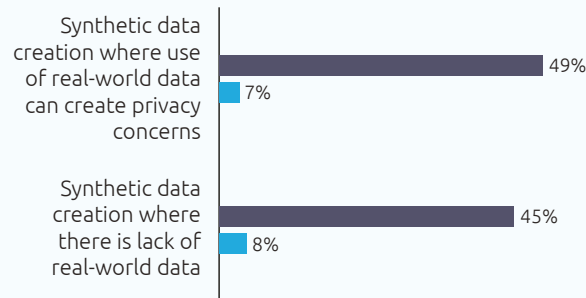
There are a variety of real-world examples of how companies use generative AI for data management on one hand and reap business benefits such as improved customer service and operational efficiency on the other.

**Figure 12.**

Organizations are using generative AI across the data value chain



### Data Creation

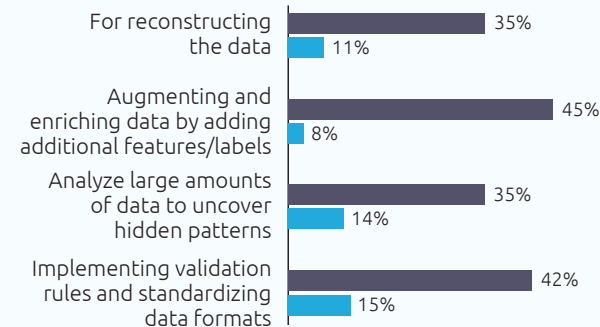


**Amazon:** Uses generative AI to generate a synthetic customer behavior data to train its computer vision system for its Just Walk Out, the system powering the Amazon Go cashier-less stores. These synthetic images and video clips mimic realistic, and sometimes rare, shopping scenarios and shoppers.<sup>1</sup>

■ Already using    ■ Planning to use in next 12 months



### Data cleansing and enrichment



**Bloomberg:** BloombergGPT, integrated into the Bloomberg Terminal, performs a range of labour-saving data cleansing functions, like swapping in ticker symbol for company names.<sup>2</sup>

■ Already using    ■ Planning to use in next 12 months

Source: Capgemini Research Institute, Data Powered Enterprise, Data Executives Survey, April 2024, N = 500 organizations represented by 500 data executives.

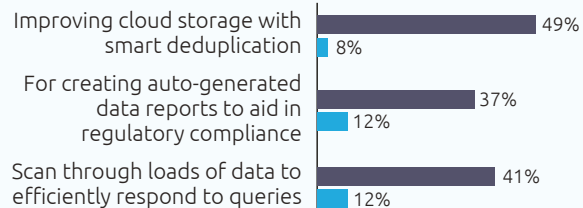
<sup>1</sup> Amazon, "How generative AI helps Amazon eliminate checkout lines and revolutionize the shopping experience," September 2023.

<sup>2</sup> Bloomberg, "Introducing BloombergGPT, Bloomberg's 50-billion parameter large language model, purpose-built from scratch for finance," March 2023.

<sup>3</sup> Forbes, "How Morgan Stanley is training GPT to help financial advisors," March 2023.



### Data curation, storage, and querying



**Morgan Stanley:** is experimenting with generative AI(GPT-4) to extract information from over 100,000 of its documents to answer queries from around 16,000 financial advisors on a variety of topics including investment recommendations, general business questions and process related questions.<sup>3</sup>

**Goldman Sachs:** is running several PoCs using the technology for document categorization, precise information extraction and sophisticated coding tasks.<sup>4</sup>

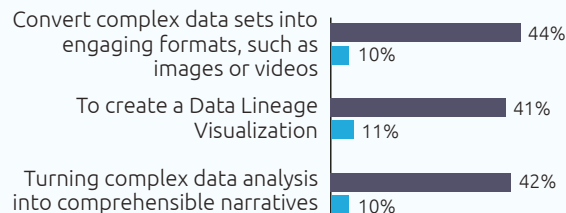
■ Already using    ■ Planning to use in next 12 months

4 The Wall Street Journal, “Goldman Sachs CIO tests generative AI,” May 2023.

5 OCBC, “OCBC is first Singapore bank to roll out generative AI chatbot to all employees globally,” October 2023.



### Data visualization



**OCBC:** has deployed several generative AI productivity tools to summarize documents, transcribe calls and create an internal knowledge base. OCBC Whisper, its speech-to-text technology, is used to transcribe and summarize calls in real time at OCBC contact centre.<sup>5</sup>

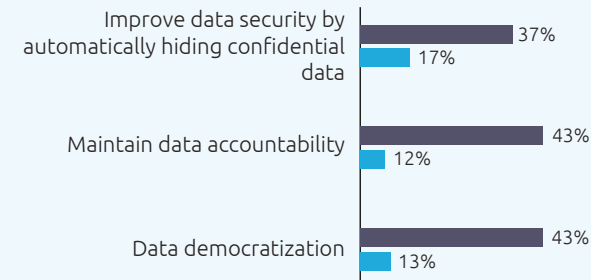
**Amazon:** Amazon’s reviews summary feature uses generative AI to provide customer with a short paragraph on the product detail page and the most frequently mentioned customer sentiment across written reviews.<sup>6</sup>

■ Already using    ■ Planning to use in next 12 months

6 Amazon, “8 ways Amazon is using generative AI to make life easier, from a more conversational Alexa to a better reviews experience,” December 2023.



### Data accountability and security



**Mastercard:** Brighterion, a Mastercard company, uses AI models can identify customers’ card behaviour and rise alerts on potential scams.<sup>7</sup>

■ Already using    ■ Planning to use in next 12 months

7 Mastercard, “AI is growing fast. How we’re growing with it,” December 2023.

04

**What benefits do data masters receive?**



Data masters constitute 17% of the organizations we surveyed and reap higher benefits across data operations, generative AI, business, and financial metrics.

## Data masters get more benefits from generative AI

Data masters derive more benefits by using generative AI. We found percentage of data executives' time saved annually by generative AI application usage is much higher at 69% for data masters compared to 55% of others. Data masters also observe improved speed in the work of data executives with 78% increase in speed of data executives' work after employing generative AI solutions compared to 70% in others.

## Data operations benefits accrued by data masters

Data masters fare better than their counterparts on multiple metrics, including quality, security, and monetization. In our 2020 research on data-powered enterprises we found among the organizations where business is trusted 77% say business teams work with

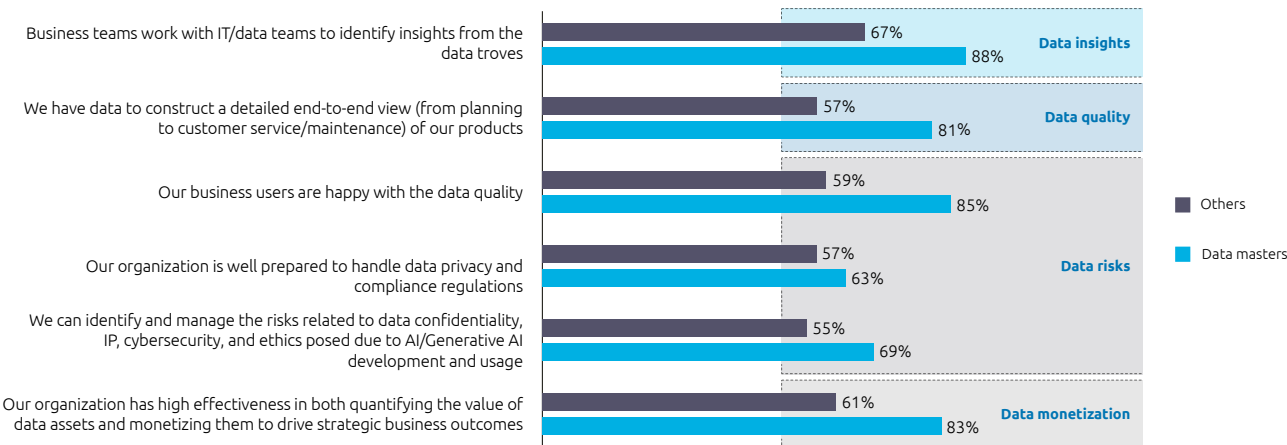
technology teams to identify insights from data troves.<sup>24</sup> This theme continues this year with 88% of data masters agreeing with this compared to 67% of others.

We also found that 83% of data masters have high effectiveness in quantifying the value of data assets and monetizing them, compared with 61% of others.

**Figure 13.**

Data operations benefits accrued by data masters span the data value chain

**% respondents reaping the following data-related benefits..**



Source: Capgemini Research Institute, Data powered enterprises survey, Data Executives Survey, April 2024, N = 500 organizations represented by 500 data executives.

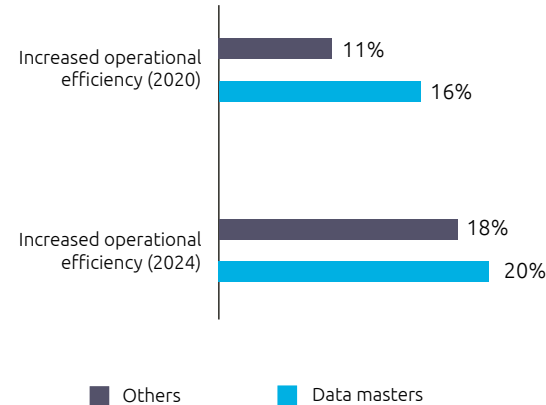
## Data masters have improved productivity with higher operational efficiency

Organizations have also reaped cost savings and efficiency improvements due to enhanced employee productivity and process efficiencies. Shipping multinational UPS is already reaping the benefits of using generative AI in its contact centers. During pilot testing, UPS earned a 50% reduction in the time agents spent resolving emails. UPS plans to apply generative AI to other functions within the enterprise, including human resources, sales, and finance.<sup>25</sup> Data masters have seen a larger reduction in customer engagement metrics such as customer complaints and have also witnessed an increase in sales due to personalized and targeted marketing enabled by data-generated insights, as well as by monetizing data and launching new products and services.

**Figure 14.**

Data masters see greater improvement in business operational metrics

### % improvement in operational efficiency through data use



Source: Capgemini Research Institute, Data Powered Enterprise, Business Executives Survey, April 2024, N = 500 organizations represented by 500 business executives.

## Benefits – financial metrics analysis

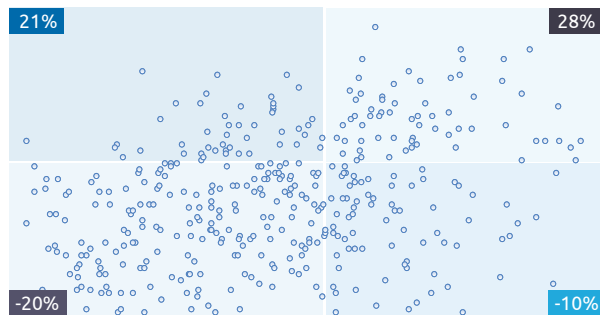
On average, data masters outperform other cohorts on multiple financial metrics due to improved operational efficiency. We noted that data masters are already showing improvement in some key metrics, such as earnings before interest and taxes (EBIT) growth, net income growth, and net income margin.

### EBIT growth over the past three years

Figure 15 indicates the difference in performance of that particular cohort from the average value for all organizations. For instance, the data masters have seen a 28% higher EBIT growth over the past three years as compared with the average, while the data laggards have realized a 20% lower growth than average. Overall, organizations that focus more on behaviors (culture, change management, leadership) and foundations together reap the most benefits.

**Figure 15.**

Data masters show a higher EBIT growth vs. other cohorts



Source: Capgemini Research Institute, Data Powered Enterprise, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.

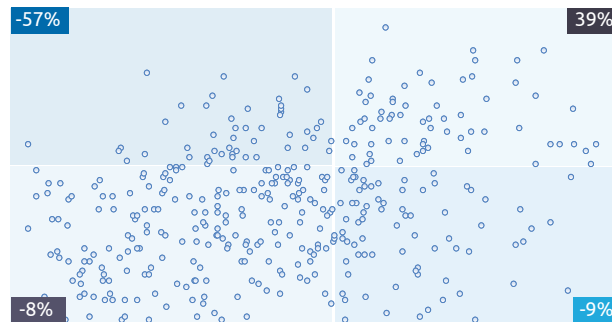
Note: We calculated the average EBIT growth over the past three years for the 500 organizations we surveyed. Then we identified the average EBIT growth for each of the four cohorts separately. Then we calculated by how much does each cohort deviate (perform better or worse) from the overall average, which is shown as the percentage in the charts below.

## Net income margin

Data masters showed a 39% higher net income margin the previous year than average. This means that this cohort's average net margin is 39% higher than the average net margin of the 500 organizations.

**Figure 16.**

Data masters show better net income margin than other cohorts



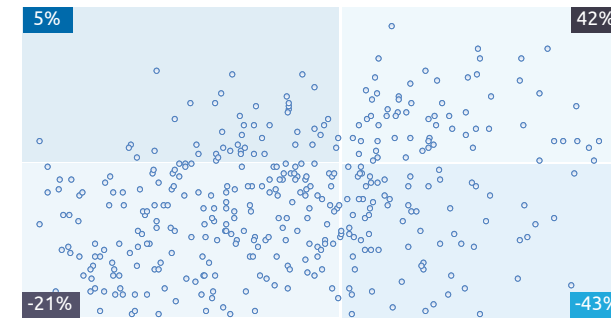
Source: Capgemini Research Institute, Data Powered Enterprise, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.

## Net income growth over three years

Data masters realized 42% higher net income growth than the average over the past three years.

**Figure 17.**

Data masters show a higher net income growth vs. other cohorts



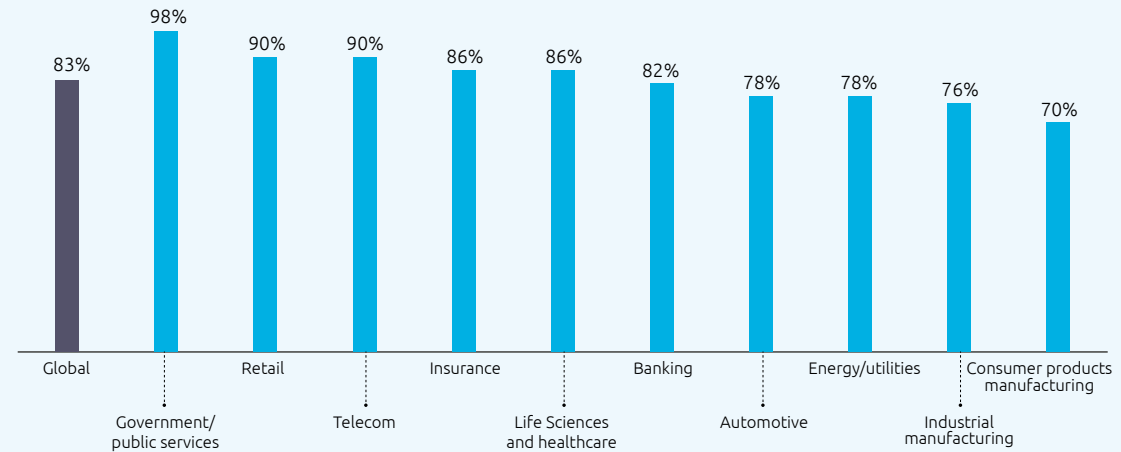
Source: Capgemini Research Institute, Data Powered Enterprise, April 2024, N = 500 organizations represented by 500 data executives and 500 business executives.

## Organizations place a high priority on sustainable data operations

Sustainability commitments and reporting obligations require organizations to view data operations through a sustainability lens. As many as 83% of business executives say ensuring sustainable data production, storage, and usage are key elements of their data strategy (see Figure 18), and 72% state they have set goals to reduce data carbon footprint. The EU Corporate Sustainability Reporting Directive addresses the disparity between investors' information needs and the available corporate sustainability information, with the first reporting cycle due in 2025.<sup>26</sup>

**Figure 18.**  
Sustainable data strategy across sectors

**Share of business executives who say ensuring sustainable data production, storage, and usage is a key component of their data strategy**



Source: Capgemini Research Institute, Data Powered Enterprise, Business Executives Survey, April 2024, N = 500 organizations represented by 500 business executives.

Organizations are striving to ensure sustainable data operations. Around 79% of business executives say they evaluate vendors (e.g. cloud solution vendors) on their sustainability credentials. Dell's CloudIQ, a cloud-based app for monitoring Dell infrastructure, tracks energy use and carbon footprint, which helps customers improve workload efficiency, consolidate resources, and upgrade to energy-efficient technology.<sup>27</sup> The Group Head of Data Management at a European bank adds: *"There's been a push over the last year or two to deal with sustainability challenges by cleaning up old data and leaving a lot of legacy data. All of those things have bled into the thinking about how we make data services and platforms sustainable and effective from an energy and cost perspective."*

**Data executives are also taking steps to reduce the environmental footprint of data:**

60%

are employing practices that seek to reduce data collection to the minimum required

54%

are optimizing data-storage architecture

43%

are optimizing data software and algorithm

Also, with generative AI becoming a forefront technology and moving into core business processes, the need to address sustainability has become more prominent. Organizations acknowledge the potential for generative AI to have a higher carbon footprint in comparison to traditional IT programs. Our previous research showed that worldwide, 78 percent of executives are aware of the greater carbon footprint from generative AI. Hence it is important to Integrate sustainability considerations into generative AI strategies and technology choices so that enterprises align well with responsible corporate practices and societal expectations.



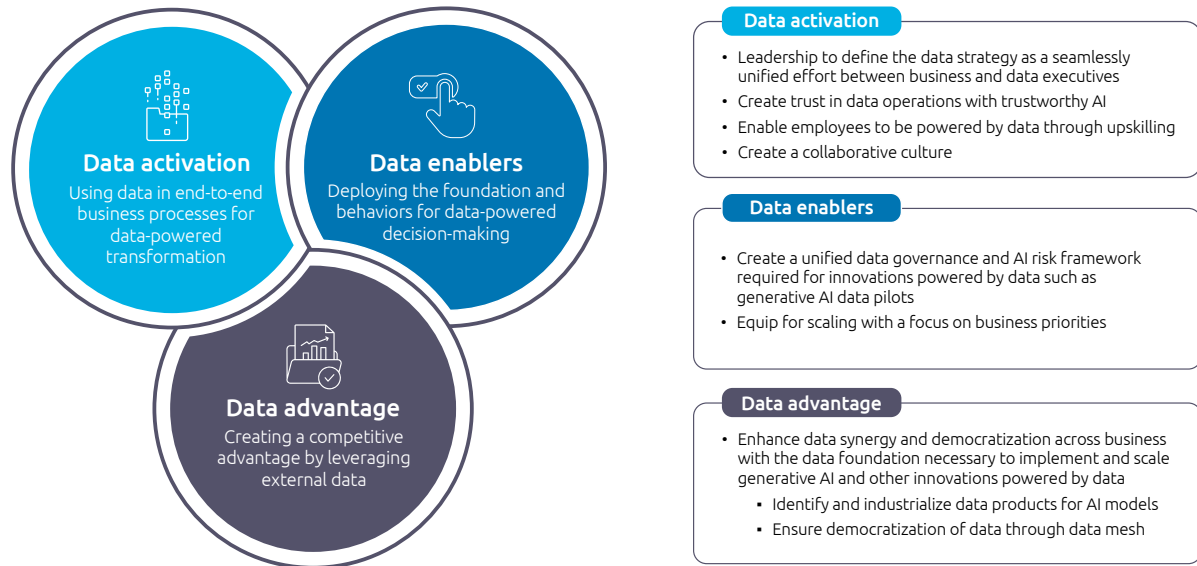
05

How are data masters leading the data transformation?

Based on our survey and interactions with industry experts, we have identified how data masters (see Figure 19) are leading the journey to become data-powered enterprises. The bar to become a data master has become higher as there are more capabilities needed to become one.

**Figure 19.**

Data masters excel across multiple initiatives



Source: Capgemini Research Institute analysis.

## Leadership to define the data strategy as a seamlessly unified effort between business and data executives

### Define the data strategy with the vision for AI/generative AI

A majority (80%) of data executives in the data masters category have defined a strategy to become a data-powered organization, compared with 61% of others. Also, 81% of data executives in the data masters category state that all business areas in their organization have a defined data/analytics strategy and roadmap, compared with 64% of others. An even larger proportion (87%) of data executives in the masters cohort have laid out a corporate vision of a data-powered enterprise with AI/generative AI at its center, compared with 62% of others. This is coupled with the fact that 84% of data executives in the data masters category state their organizational strategy and vision on AI/generative AI are readily accessible to their employees, compared with 64% of others.

### Ensure synergy between data strategy and overall organizational strategy

A disconnected data strategy leads to misdirected data teams, leading, in turn, to a disconnect with the business teams. Mónica León Santamaría, Global Data and AI Center of Enablement Lead at Bupa, underlines the importance of data strategy alignment: *“Data strategy must be dictated by business strategy across three key aspects: data culture and literacy, data governance, and data management.”* Nearly four out of five (79%) data executives in the data masters category have a centralized IT/business team that defines data strategy, draws up a

roadmap, and governs the data-powered decision-making and implementation, compared with 64% of others.

# 80%

of data executives in the data masters category have defined a strategy to become a data-powered organization





### Create a C-suite data-leadership position

A C-suite data leader ensures accountability for data initiatives and close alignment with business needs. A high 94% of data masters' chief data officers (CDOs) report to the CEO, compared with 79% of others.

C-level accountability for data initiatives also ensures organizations to operationalize initiatives into digestible building blocks, closely following the business capabilities through the form of data domains. Through the domains, the CDO orchestrates initiatives close to the business problems and collaborates with the data experts, who understand each domain's data, to generate high-quality and reusable data products.

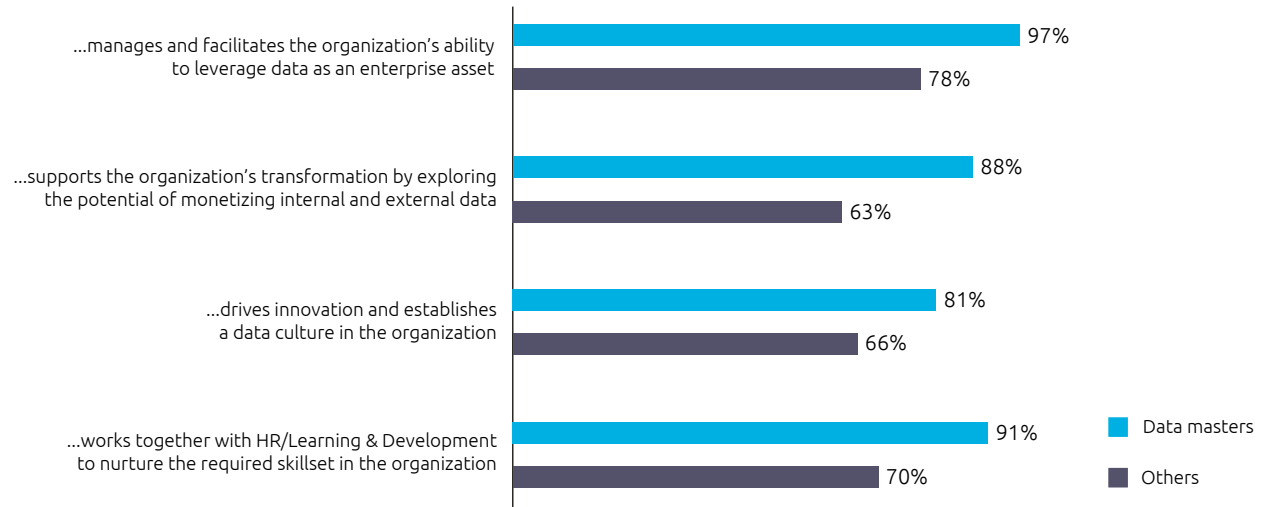
### The strategy must assign business ownership and democratize data through data domains

A lack of ownership of data on the business side continues to be a key issue for organizations on their journey toward becoming data-driven. While only 41% of business executives in the data masters category are held back by a lack of ownership on the business side, 57% of others state that a lack of business ownership is a major hurdle to becoming data-driven. The key is to establish data domains following the business processes, capabilities, and architecture.

**Figure 20.**

Data masters exhibit stronger data leadership

#### Share of executives who say their CDO...



Source: Capgemini Research Institute, Data powered enterprises survey, Data Executives Survey, April 2024, N = 500 organizations represented by 500 data executives; N = 86 for data masters and N = 414 for others.

### Involve senior leadership in the decision-making process

Senior leadership support for data projects is essential. Over half (56%) of data masters say key business stakeholders promote generative AI, compared with 44% of others. A high 90% of data masters also cite synergy between the data team and overall leadership in implementing generative AI solutions, compared with 71% of others.

# 90%

of data masters cite synergy between the data team and overall leadership in implementing generative AI solutions

## Enhance data synergy and democratization across the business with the data foundation necessary to implement and scale generative AI and other innovations powered by data

### Identify and industrialize data products for AI models

As data scientists build more AI-based data models, a key challenge they face is that less than 20% of these models go into production deployment.<sup>28</sup> Jaydeep Ghosh, Data Management and Analytics professional in healthcare industry, adds: *“Organizations want to merge their internal data with the power of generative AI to create more specific insights. We have to define the good data sources.”* As much as 87% of data executives in the data masters category have set standards and prequalify sources to

ensure the data that is fed into the AI/generative AI models follows specific protocols, compared with 69% of other organizations.

Organizations are using data products such as reusable datasets to solve business problems. To have meaningful impact, this must be built at organizational level and reflect the business problems of the data domain. Assigning ownership of data products at this level ensures that the developed data products are aligned with the demand within their ecosystem. Three-quarters (76%) of data masters are working on data modernization to ensure alignment of data within the organization and wider ecosystem to scale generative AI deployment, compared with 65% of others.

### Ensure democratization of data through data mesh

Democratization of data is essential to providing the required data to the business users at speed and in the right form and granularity. This enables business users across the organization to explore data and derive actionable insights without waiting for the technical teams. Data mesh enables this through data products and their metadata being available in a single platform for different business units, domain-oriented decentralized data ownership, self-serve data infrastructure as a platform, and federated

computational governance. This has multiple advantages, according to a CDO: *"Making data available to all of the business areas will ensure that the AI that you build is fast, efficient, and driven by the right data consistently across different teams."*

A strong 79% of business executives in the data masters category state that generative AI will require significant changes to various aspects of data management, including data collection, storage, retrieval, and governance, compared with 61% of others. Four out of five business executives in the data masters category say they are architecting their applications to be data-powered, compared with 64% of others.

Making data available in a reusable format, especially when multiple teams need to use the data, is vital. Swapna Tom, Distinguished Technologist – Enterprise Data and Analytics, HP mentions on vector databases: *"We do the indexing and vectorization once, creating a common layer so that the same document need not be indexed, even if you have two or three use cases. A good example is a gen-AI virtual assistant. If you have multiple use cases, we reuse that data, rather than everybody trying to reinvent the wheel."* An estimated 83% of data executives in the data masters category have evaluated the current state of data infrastructure, systems and processes, and have a comprehensive plan for scaling the deployment of generative AI, compared with 56% of others.



**"We do the indexing and vectorization once, creating a common layer so that the same document need not be indexed, even if you have two or three use cases. A good example is a gen-AI virtual assistant. If you have multiple use cases, we reuse that data, rather than everybody trying to reinvent the wheel."**

#### **Swapna Tom**

Distinguished Technologist -  
Enterprise Data and Analytics, HP

## Create a unified data governance and AI risk framework required for innovations powered by data such as generative AI data pilots

### Create a generative AI council to make informed decisions on requests

Organizations are inundated with requests for generative AI-based requests for data pilots. A governance council is required to scrutinize requests against multiple parameters and make a decision. It is vital to ensure that this council is involved throughout the entire AI life cycle. Already, 77% of data executives in the data masters category say their data teams monitor and guide all business teams working on data for ethical issues, compared with 61% of others.

### Create a framework to approve or reject generative AI/AI pilots for data operations

Multiple parameters must be evaluated before deciding on

# 77%

of data executives in the data masters category say their data teams monitor and guide all business teams working on data for ethical issues

cost, timelines, and quality of data. Demand management should also include regulatory and compliance checks before the development of use cases is triggered. A senior data executive adds: *“A value framework to evaluate generative AI use cases needs to look at ROI, complexity, data and technology availability, timelines, and integration licensing. Quality of data and how structured or unstructured it is, and how much preparation of data needs to be done, are additional parameters that we evaluate.”*

### Develop guardrails around generative AI data applications

It is important not to connect generative AI data use cases across the organization to ensure a consistent approach so that they are scalable. A set of guardrails is essential so that the business and data teams are aware of where each

proposal is in the assessment process. As many as 87% of data executives in the data masters category utilize technical guardrails and human-in-the-loop supported controls, compared with 59% of others. For instance, organizations need to protect themselves against prompt injection attacks where a hacker’s prompt disguised as benign user input is fed to an LLM application to override the LLM’s system instructions to make the app into the attacker’s tool.

### Measure the cost and the benefits to quantify the investment

Having implemented the pilot, measuring the cost and the benefits is also important to decide whether or not to scale it to production. Swapna Tom from HP confirms: *“Both cost control and innovation are required. If you don’t invest, you won’t be able to innovate.”*

A full 90% of data executives in the data masters category state that key business and technical stakeholders understand commercial models and the cost implications of generative AI technologies, compared with 73% of others. Around 76% of data masters are aware of the cost sensitivity of implementing generative AI in data operations, compared with 62% of others. It is important to take an incremental approach to get leadership buy-in and the confidence to scale such initiatives.

# Create trust in data operations with trustworthy AI

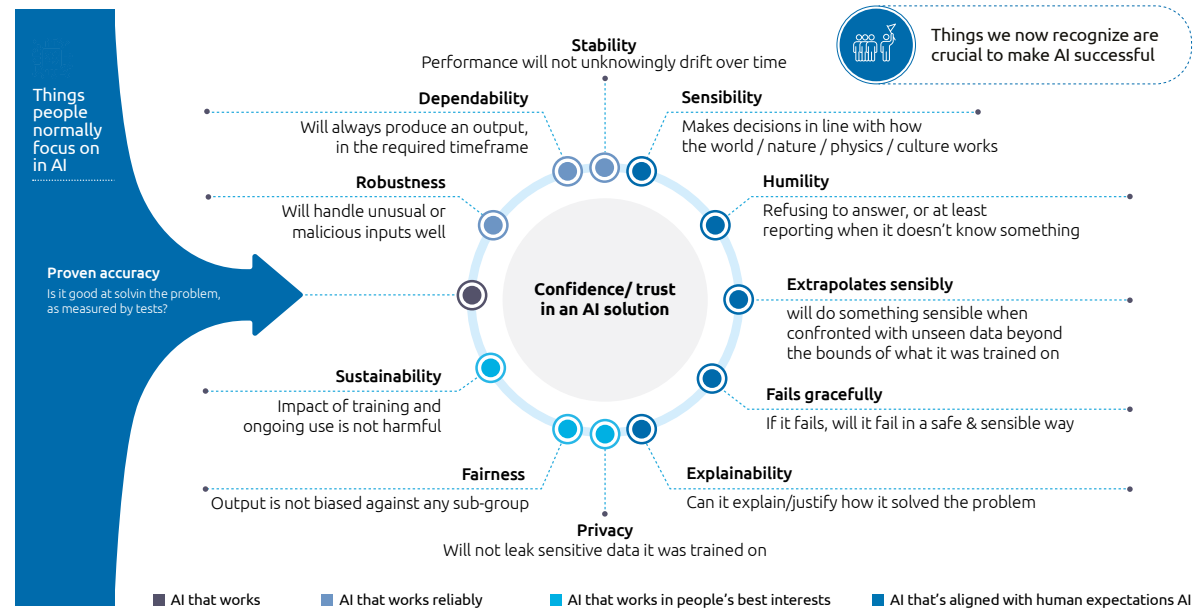
Confident adoption of AI relies not just on creating AI that works, but on creating AI that works reliably, is aligned to human expectations, and works in people's best interests. Teams involved in AI systems design and implementation, would need to think beyond accuracy, and adopt a more multi-faceted approach including transparency, ethics, and fairness.

# 91%

data masters have ethical and human-impact assessment processes to ensure that generative AI-based tools do not negatively impact data consumption, human creativity, or judgement

**Figure 21.**

In addition to model and input data accuracy, there are some other equally critical success factors for creating a trustworthy AI



Source: Unleashing Confidence in AI, Capgemini, March 2024.

**Data masters have implemented multiple steps to enable confident adoption of AI and to create trust in data operations:**

**Governance**

- 84% of data executives in the data masters category review the generative AI vendor’s data sourcing, management, security, and standards, compared with 61% of others and 91% have ethical and human-impact assessment processes to ensure that generative AI-based tools do not negatively impact data consumption, human creativity, or judgment, compared with 65% of others.
- 72% of business executives in the data masters category say they have documented policies on the source, usage, access, and processing of data in generative AI, compared with 60% of others, and 78% have a clear roadmap and strategy to address legal complications arising from underlying data used in AI/generative AI applications, compared with 68% of others

■ 77% of data executives in the data masters category say their users understand the regulatory and professional responsibilities and liabilities in the context of generative AI compared with 57% of others.

■ 57% of business executives in the data masters category ensure they use AI solutions with delimited impact, fairness, sustainability, transparency and explainability, controllability, robustness, and respect for privacy, compared with 43% of others.

**77%**

of data executives in the data masters category say their users understand the regulatory and professional responsibilities and liabilities in the context of generative AI



**“Every employee must recognize the importance of data management and data governance to ensure accurate use of the data and trust in the data for decision making.”**

**Mónica León Santamaría**  
Global Data and AI Center of Enablement Lead, Bupa

## Bias

- 83% of business executives in the data masters category are setting up processes (e.g., using algorithms) to detect bias in data used for generative AI tools (e.g., using algorithms that are specifically designed to detect bias in generative AI tools), compared with 63% of others and 67% are able to validate data generated by generative AI tools, compared with 53% of others.
- 63% of data executives in the data masters category regularly review AI/generative AI outputs for accuracy and factual correctness, including mitigation of biases and fact-checking, compared with 44% of others.

## Data confidentiality and cybersecurity

- 77% of data executives at masters take care not to feed confidential information into generative AI tools, especially if they do not have direct control over development and deployment, compared with 65% of others and 87 have defined a strategy to prevent unintentional 'leakage' of sensitive data to public LLMs, compared with 71% of others.
- 69% of data executives in the data masters category identify and manage risks related to data confidentiality, IP, cybersecurity, and ethics posed by AI/generative AI development and usage, compared with 55% of others.

Additionally, adopting robust evaluation metrics and benchmarking are some other effective ways in which organizations can build trust in generative AI models and output. Evaluation metrics like Recall-Oriented Understudy for Gisting Evaluation (ROUGE) and Bilingual Evaluation Understudy (BLEU) and benchmarking frameworks like "Big-Bench" Beyond the Imitation Game benchmark, Holistic Evaluation of Language Model (HELM) can be instrumental in instilling confidence in Gen AI output.

### Adhere to data-quality governance

Creating a data-quality team to filter and verify sources is important. The data quality team must work closely with the data creators and assure data stewardship is assigned and enforced close to its origin, as the quality cannot drastically be improved after several stages of transformation. Jaydeep Ghosh, Data Management and Analytics professional in healthcare industry, confirms: *"Good data governance is key to utilizing data assets effectively, whether for operations, business intelligence, or generative AI. We should have visibility of where the data is created. How current is it? Is it good quality? Who is the owner? What are the different data assets?"* A good 80% of data executives in the data masters category have clear protocols and guidelines for the use of data-scraping tools for preparing data to train generative AI tools, compared with 57% of others.



## Enable employees to be powered by data through upskilling

Fast-paced development of the data ecosystem has led to the need for rapid upskilling. A senior data executive confirms: *“Data skills only last two years because the tools and technology keep changing. What the business expects from a data engineer, visualization expert, or data scientist is a constant evolution of skill sets. The availability of talent in this area has begun to catch up with demand. But I still see a shortage of top-tier people.”* Around 71% of business executives in the data masters category say they are actively upskilling the workforce for the required skills in generative AI, compared with 56% of others.

Organizations have also started to upskill workers outside data teams. Mónica León Santamaría at Bupa agrees: *“We are developing a data-literacy program to ensure employees can make their own data-based decisions. This is challenging, because the required level varies. For instance, a data scientist must have a deeper understanding of the analytics techniques that they have to apply when building data models. But a business analyst must know how to use the data assets to understand the business situations.”* As many as 87% of data executives in the data masters category are actively training employees to recognize the risks of using generative AI tools, including hallucinations and biases, compared with 70% of others.

# 71%

of business executives in the data masters category say they are actively upskilling the workforce for the required skills in generative AI



**“The best way to scale generative AI data initiatives is via business teams. A business view is also essential to decide on adoption and scaling strategies. That works much better than where you've got lots of different business units trying to have a go at it.”**

**Martin Stanley**  
Imperial Brands



## Create a collaborative culture

A cultural change is required to shift to a more data-centric mindset. Employee awareness of data management and operations is essential. After all, it is usually employees within operational teams that create data records through business processes, that have the highest lever to impact data quality. Mónica León Santamaría at Bupa adds: *“Every employee must recognize the importance of data management and data governance to ensure accurate use of the data and trust in the data for decision making.”* Over half (55%) of data executives in the data masters category have high maturity of non-technical foundations in the use of generative AI, compared with 37% of others.

Collaboration with data teams across the organization facilitates cross-pollination of ideas. Mónica at Bupa adds: *“We also developed many big events across the company. We recently hosted an international data summit, where different teams shared their data initiatives. We bring external providers to explain their initiatives and host hackathons, training programs, workshops, and many other events to enhance data culture. We need people to think that the data is something to come and enjoy.”* A high 95% of data executives in the data masters category build cross-functional data and insights teams

to work with business sponsors, business analysts, data engineers, data scientists, solution architects, and software developers, compared with 69% of others. These teams are typically organized around data domains. Also, 66% say their data governance makes it easier for teams from different business units to collaborate, compared with 44% of others.

## Equip for scaling with a focus on business priorities

### Scale via business teams

A business-led team for scaling initiatives ensures visibility and traction with business priorities. Martin Stanley of Imperial Brands adds: *“The best way to scale generative AI data initiatives is via business teams. A business view is also essential to decide on adoption and scaling strategies. That works much better than where you’ve got lots of different business units trying to have a go at it.”* Around 78% of data executives in the data masters category say the budget for generative AI data initiatives is allocated centrally, compared with 66% of others. The budget for platform engineering, governance, creating guardrails and communications are allocated and

utilized centrally. The tools are made available to the business teams for scaling the identified use cases. HP’s Swapna Tom confirms: *“We have a gated process. Gate one is about pilots. Anyone can do it and mostly we enable it in a controlled environment. At gate two, when you’re going to production, we look into all governance criteria such as cyber, legal, and privacy. So, those two gates help us to manage this in a better way and there is no slowdown or priority conflict because it’s a joint decision.”*

### Keep overall business priorities in mind when deciding on scaling

A key criterion to consider before scaling initiatives is whether the initiative affords higher visibility to your business priorities and encourages management buy-in. This also ensures the organization utilizes data resources for key priorities. While 87% of business executives in the data masters category have a clear strategy and capability to scale generative AI implementation, compared with 77% of others. Data masters are also aware of key use cases to scale. Three-quarters (76%) also say they can target the top generative AI use cases to be unlocked in production at scale compared with 63% of others.

# Conclusion

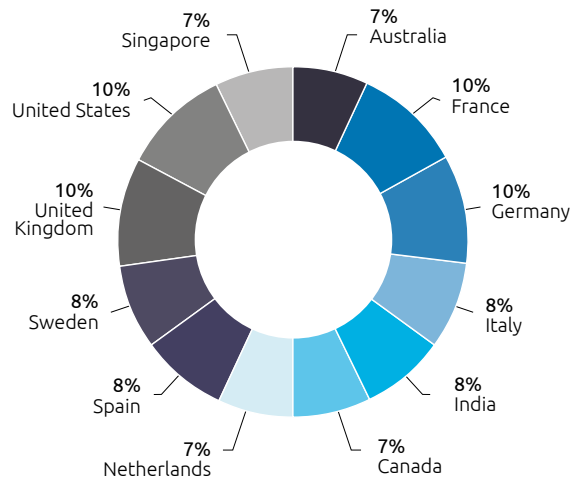
Compared with a few years ago, today organizations overall are more effective at embedding data and insights into core business processes and using proprietary data to drive business goals. As a result, the percentage of data laggards has come down from previous years. However, to become data masters and fully leverage the potential of data, organizations need to up their game. Unstructured data, the rise of generative AI, and growing ethical, legal, and security vulnerabilities put new demands on organizations, which need to be suitably addressed in order to become fully data-powered. Hence, in order to become data masters, organizations would need to cover substantial ground on multiple dimensions, including developing new-age data skills, establishing robust data governance, and business-data collaboration. Having said that, the evident benefits around operational and financial performance that come with data mastery make it worth the effort.

# Research methodology

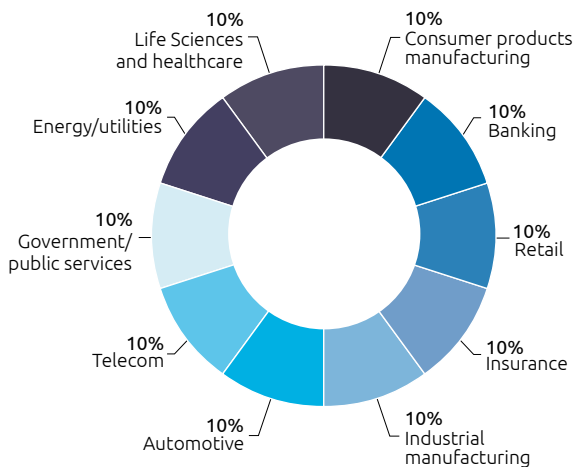
The global survey took place in April 2024. We surveyed 500 business executives and 500 data executives employed at 500 organizations, each with more than \$1 billion in annual revenue, across 12 countries and 10 industries in North America, Europe, and Asia-Pacific. Executives surveyed were director-level and above and from business functions including strategy, operations, sales, and marketing, and data

functions, such as IT, AI and analytics, data management, architecture, governance, and information security. We surveyed two executives from every organization. The distribution of executives and their organizations is provided in the following figures. We also interviewed 10 senior executives for their perspectives on this topic.

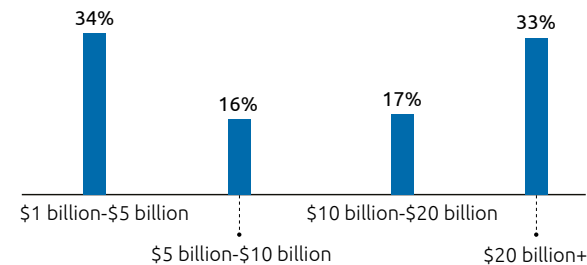
### Organizations by country



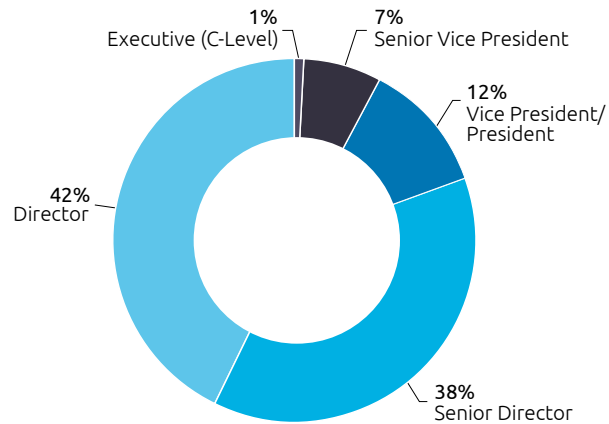
### Organizations by sector



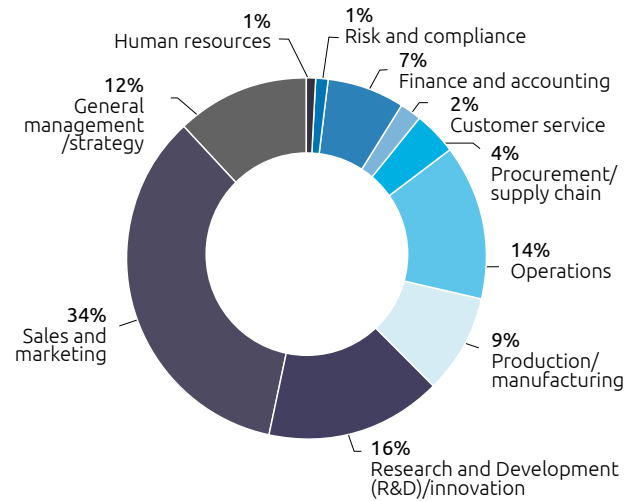
### Organizations by annual revenue



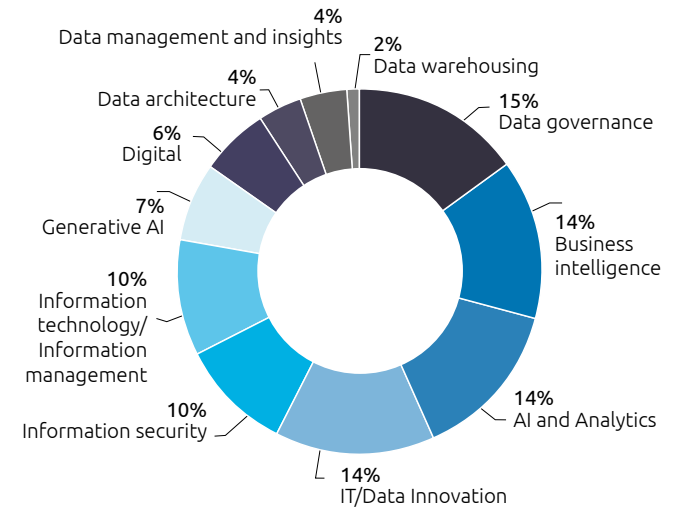
### Respondents by job title



### Business executives by function



### Data executives by function



# Appendix

## Statements considered for the technology executives' model

| Category  | Theme                                   | Statement  |
|---|---|--|
| <ul style="list-style-type: none"> <li>Data Foundations (tech and tools)</li> </ul> | <b>Identify all its data and models</b> | We have data to construct a detailed end-to-end view of our product (from planning to customer service/ maintenance), such as a digital twin   |
|   |   | Our organization has a complete data inventory   |
|   |   | We have a record of the data owner(s) for all internal data  |
|   |   | We have the required data to train generative AI models  |
|   |   | We use external data from suppliers, distributors/retailers; platform providers; anonymous consumer data (cookies, etc.); personally identifiable information (PII) of consumers; consumer usage data; social media listening data; data from blogs/product reviews, etc.; publicly available competitor data; analyst/Industry reports; data from hyperscalers such as Google, Amazon, Facebook, etc.; proprietary datasets from data aggregators such as Nielsen, Experian, etc.; open data (government and other public-sector data) for data-powered decision making and actioning |
|   | <b>Create and collect data</b>          | We create data by designing products and processes to capture new data   |

| Category  | Theme   | Statement   |
|---|---|---|
| <ul style="list-style-type: none"> <li>▪ Data foundations (tech and tools)</li> </ul> | <p><b>Process and harvest data and models</b></p> | <p>We are able to use structured, semi-structured, and unstructured data for data-powered decision-making and actioning</p>   |
|   |   | <p>Data systems and capabilities, based on their readiness to implement generative AI applications, are highly mature</p>   |
|   |   | <p>We have defined and continue to update an enterprise-wide data catalogue to assist data users</p>  |
|   |   | <p>We have automated data collection</p>  |
|   |   | <p>We can store, retrieve, and analyze the data across the value chain</p>  |
|   |   | <p>We have developed data-management processes that promptly source, clean, prepare, integrate, and provide access to data at the speed that the business needs</p> |
|   |   | <p>We have deployed processes to check quality of data at points of capture</p>   |
|   |   | <p>We ensure data stewardship in each data domain to improve data quality</p>   |

| Category  | Theme  | Statement  |
|---|--|--|
| <ul style="list-style-type: none"> <li>■ Data foundations (tech and tools)</li> </ul> | <b>Data governance implementation</b>            | Our organization is aware that strict data governance is the basis for successful use of generative AI   |
|   |  | Our data governance considers and plans for varying maturity levels across business units  |
|   | <b>Scale infrastructure, platform, and tools</b> | We are expanding data, BI and analytics in the cloud   |
|   |  | We have evaluated the current state of data infrastructure, systems and processes and have a comprehensive plan for data modernization for scaling deployment of generative AI |
|   |  | Share of our organization's data sources in cloud  |
|   |  | We have invested in analytics tools and platforms  |
|   |  | We are offering data-preparation tools for self-service data management  |
|   |  | We are offering data-preparation tools for self-service data management  |
|   | <b>Data and AI platform</b>                      | We have dedicated resources for retraining and redeployment of existing models -Data and AI platform   |

| Category   | Theme                          | Statement  |
|--|--------------------------------|--|
| <ul style="list-style-type: none"> <li>Data behaviors</li> </ul> | <b>Activate</b>                | We use descriptive, diagnostic, predictive, prescriptive, autonomous/self-optimizing/generative approaches for decision-making in our organization |
|  |                                | We continually act on insights to optimize business outcomes   |
|  | <b>Foster culture</b>          | Business teams work with IT/data teams to identify insights from the data troves   |
|  |                                | We have invested in our data culture by enabling employees with the skills and tools to generate and apply insights                                |
|  |                                | We actively promote exploration, collaboration on new ideas and experimentation at all levels  |
|  |                                | Employees can work on new ideas/prototypes without worrying about failure  |
|  | <b>Data guiding Principles</b> | Data is FAIR, secured, and sustainable   |
|  |                                | We have updated our policies related to IT, AI, confidentiality, and data governance to meet generative AI challenges                              |
|  |                                | Executive leaders champion data privacy, security and ethics, and are accountable for performance  |



| Category   | Theme                                      | Statement   |
|--|--|---|
| <ul style="list-style-type: none"> <li>▪ Data behaviors</li> </ul> | <b>Data activation vision and strategy</b> | We upskill our employees on data skills such as model training, course correction, and maintenance, in addition to their regular AI/ML skillset/generative AI |
|  |  | We have role-based data-upskilling programs for most of our employees   |
|  |  | We have created and filled in senior positions for data-policy governance for generative AI   |
|  |  | We train our business users on analytical and storytelling skills   |
|  | <b>Foster data culture</b>                 | Our senior leadership is fully committed to appropriate investments in resources and technology towards becoming data-powered                                 |
|  |  | Our data strategy includes AI/generative AI-related strategy, and metrics for delivering business impact  |
|  |  | Our business strategy communicates how we will use data to drive strategic decisions, business outcomes, and to create competitive differentiation            |
|  |  | Our data/analytics officer ensures that the data/analytics strategy is aligned with overall business strategy   |
|  |  | We build cross-functional data and insights teams that work with data engineers, data scientists, solution architects and software developers                 |

## Statements considered for the business executives' model

| Category  | Theme                                      | Statement  |
|---|--|--|
| <ul style="list-style-type: none"> <li>Data Enablers</li> </ul> | <b>Identify all its data and models</b>    | We have data to construct a detailed end-to-end view (from planning to customer service/maintenance), such as a digital twin, of our products  |
|   |  | We have a good understanding of the data and its provenance  |
|   |  | We use external data from suppliers, distributor/retailer, platform providers, and anonymous consumer data (example cookies etc.), personal Identifiable Information (PII) of consumers, consumer usage data, social media listening data, data from blogs/product reviews etc., publicly available competitor data, analyst/Industry reports, data from hyperscalers like Google, Amazon, Facebook etc., proprietary datasets from data aggregators like Nielsen, Experian etc., open data (Govt and other public sector data) for data powered decision making and actioning |
|   | <b>Create and collect data</b>             | We use generative AI for generating new data (or synthetic data) to augment existing data resources  |
|   |  | Our products and processes are designed to capture new data  |
|   | <b>Process and harvest data and models</b> | We are easily able to access and analyze multiple data types for analysis (e.g., structured, semi-structured, unstructured, customer, web analytics, and advertising data)   |

| Category  | Theme                                      | Statement  |
|---|--|--|
| <ul style="list-style-type: none"> <li>■ Data Enablers</li> </ul> | <b>Process and harvest data and models</b> | Our data catalog is clearly defined and regularly updated for us to use  |
|   |  | Our data collection process is automated   |
|   |  | Data systems and capabilities based on their readiness to implement generative AI applications are highly mature   |
|   |  | We can easily store, retrieve and analyze the data at all levels of the organization's value chain   |
|   |  | We get access to data at the speed at which we need  |
|   |  | We have access to self-service analytics   |
|   |  | We do not have sufficient data quality checks at the points of capture   |
|   |  | We have a good quality of data for each of the following data types – customer data, product data, sales data, supply chain data, finance data, operations data, employee data, external data, services data, and partner/platform provider data |

| Category        | Theme                                     | Statement   |
|-----------------|---|---|
| ■ Data Enablers | Data governance implementation            | Our organization is aware that a profound data governance is the basis for successful generative AI         |
|                 |   | Our data governance considers and plans for varying maturity levels of each business unit                   |
|                 | Scale infrastructure, platform, and tools | We have expanded data, BI and analytics in the cloud/ We have expanded data, BI, and analytics in the cloud |
|                 |   | We have invested in analytics tools and platforms   |
|                 |   | Our organization has dedicated resources for retraining and redeployment of existing models                 |

| Category  | Theme                      | Statement  |
|---|----------------------------|--|
| <ul style="list-style-type: none"> <li>■ Data Behaviors</li> </ul>  | <b>Activate data</b>       | We use predictive, prescriptive, autonomous/self-optimizing approaches for decision making in each of the following functions – general management, sales and marketing, R&D/innovation, production/manufacturing operations, supply chain and procurement, customer service, finance and accounting, risk and compliance, human resources |
|   |                            | Generative AI is on track to help us utilize our data better and help us improve our products and services.  |
|   |                            | We continually act on data-generated insights to optimize business outcomes  |
|   | <b>Foster data culture</b> | Business teams work with IT/data teams to identify insights from the data troves   |
|   |                            | We have invested in a data culture by enabling employees with the skills and tools to generate and apply insights  |
|   |                            | We actively promote the exploration, collaboration of new ideas and experimentation at all levels  |
|   |                            | Employees can work on new ideas/prototypes without worrying about failures   |
|   |                            | We build cross-functional data and insights teams that work with data engineers, data scientists, solution architects and software developers  |
|   |                            | Our business strategy communicates how we will use data to drive strategic decisions, business outcomes and to create a competitive differentiation  |
|   |                            | Our data/analytics officer ensures that the data/analytics strategy is aligned with the overall business strategy  |
| Our senior leadership is fully committed to appropriate investments in resources and technology to make us data-powered |                            |  |

| Category   | Theme                                      | Statement   |
|--|--|---|
| <ul style="list-style-type: none"> <li>■ Data Behaviors</li> </ul> | <b>Data activation vision and strategy</b> | Our employees realize the value of being data-powered   |
|  |  | We have role-based data upskilling programs for most of our employees   |
|  |  | We are trained on analytical and storytelling skills  |
|  | <b>Unlock value</b>                        | We monetize data assets/insights through our products and services  |
|  | <b>Data guiding principles</b>             | Sustainability (data production, storage and usage are sustainable) of data is a key consideration in our data governance |
|  |  | Access policies are clear and defined for all kinds of user roles   |
|  |  | We have updated our policies related to IT, AI, confidentiality, and data governance to meet the generative AI challenges |
|  |  | Executives leaders champion data privacy, security and ethics and are accountable for success                             |

# References

1. European Commission, "European data strategy," 2023
2. *Consumergoods.com*, "Coca-Cola RGM Levers: Expand AI Retail Personalization & Adjust Price Pack Architecture."
3. *Consumergoods.com*, "With Coca-Cola's RGM Strategy, An Intelligent Prioritization Model Is Key," Sept 12, 2023.
4. By 'activated data. we mean embed data and insights into the core business processes alongside AI/generative AI and enable business ownership of data to drive business goals.
5. Economist, "The future of Europe's data economy," 2022.
6. Capgemini Research Institute, "The data-powered enterprise: Why organizations must strengthen their data mastery," 2020.
7. Note: 38% of organizations agreed with the statement, "We actively promote data-driven decision-making" in 2018; 50% of organizations agreed with the statement, "Decision-making in our organization is completely data-driven" in 2020; 66% of data executives agreed with the statement, "Decision-making and actioning in our organization is completely data-driven" in 2024; 53% of business executives agreed with the statement, "We continually act on data-generated insights to optimize business outcomes" in 2024.
8. Kaiserpermanente, "The future of health care is digital," Oct 23, 2023; Forbes, "Being Data-Driven Is Likely Your Best Bet," July 18, 2023.
9. Just-food.com, "Leading consumer goods companies in the AI theme," May 22, 2023.
10. *Aidataanalytics.network*, "Coca-Cola vs. Pepsi: The Sweet Fight For Data-Driven Supremacy," 2022.
11. Note: in 2024, 50% of organizations agreed with the statement: "Decision-making in our organization is completely data-driven" in 2020; 66% of data executives agreed with the statement, "Decision-making and actioning in our organization are completely data-powered;" and 53% of business executives agreed with the statement: "We continually act on data-generated insights to optimize business outcomes."
12. Capgemini, "THE VEHICLE DATA BIG BANG How to turn theory into reality", May 2022.
13. Adweek, "The top 8 retail media networks for advertisers In 2024," April 2024.
14. Chase, "Chase launches Chase Media Solutions, a new digital media business, connecting 80 million US consumers with the brands they love," April 2024.
15. Progressive Insurance website, accessed May 2024.
16. Flyaps.com, "AI and data monetization in telecom," December 12, 2023.
17. Adweek, "T-Mobile Advertising Solutions launches Retail Media Network in NewFronts debut," April 30, 2024
18. Capgemini Invent, The vehicle data big bang: How to turn theory into reality, May 2022.
19. Uber website, accessed May 2024.
20. World Economic Forum, "Data volume is soaring. Here's how the ICT sector can sustainably handle the surge," May 2024
21. Note: In 2024, 45% of business executives globally disagreed to this statement: 'We do not have a complete picture of all the data inventory. In 2020, 25% of business executives globally disagreed to this statement: 'We do not have a complete picture of all the data inventory.

22. CIOdive.com, "How P&G rolled out its internal generative AI model," October 18, 2023
23. Ciodive.com, "Generative AI jobs, though rare, are on the rise: Indeed," March 18, 2024.
24. Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=313 business executives who agree to the statement, "we do not entirely trust the data we get."
25. CIO.com, "UPS delivers customer wins with generative AI," May 3, 2024.
26. UnitofMeasure, "Making ESG reporting a part of your data strategy," February 2024.
27. Forbes, "Enterprise Data Technology Part 4 — Sustainability Data," March 2024
28. HBR, "Why your company needs data-product managers," October 2022.





# Key contributors

**Niraj Parihar**

CEO, Insights & Data, Capgemini  
[niraj.parihar@capgemini.com](mailto:niraj.parihar@capgemini.com)

**Steve Jones**

Executive Vice President, Data Driven  
Business & Gen AI, Capgemini  
[steve.g.jones@capgemini.com](mailto:steve.g.jones@capgemini.com)

**Ron Tolido**

CTO, Insights and Data Global Business Line,  
Capgemini  
[ron.tolido@capgemini.com](mailto:ron.tolido@capgemini.com)

**Mark Oost**

Global Offer Leader Generative AI,  
Global Offer Leader AI, Analytics and Data  
Science, Insights and Data, Capgemini  
[mark.oost@capgemini.com](mailto:mark.oost@capgemini.com)

**Padmashree Shagrithaya**

EVP & MD, Insights & Data India, Capgemini  
[padmashree.shagrithaya@capgemini.com](mailto:padmashree.shagrithaya@capgemini.com)

**Aruna Pattam**

Client Partner & Generative AI SME,  
Insights & Data, Asia Pacific, Capgemini  
[aruna.pattam@capgemini.com](mailto:aruna.pattam@capgemini.com)

# Key contributors

**Jerome Buvat**

Head of the Capgemini Research Institute

[jerome.buvat@capgemini.com](mailto:jerome.buvat@capgemini.com)

**Sumit Cherian**

Program Manager, Capgemini Research Institute

[sumit.cherian@capgemini.com](mailto:sumit.cherian@capgemini.com)

**Subrahmanyam KVJ**

Senior Director, Capgemini Research Institute

[subrahmanyam.kvj@capgemini.com](mailto:subrahmanyam.kvj@capgemini.com)

**Tvishee Kumar**

Senior Manager, Capgemini Research Institute

[tvishee.kumar@capgemini.com](mailto:tvishee.kumar@capgemini.com)

The key contributors would like to especially thank Siva Chidambaram, Jeevan Joseph Mattapparampil, Prateek Gondwal, and Vaishnavee A from the Capgemini Research Institute for their contributions to this research.

The key contributors would like to thank Anne-Laure Thibaud, Andy Vickers, Alex Maradon, Amit Kumar Bhattacharya, Eric Reich, Robert Engels, Ramana Bhandaru, Bikash Dash, Eric Grumblatt, Aurobindo Saha, Manuel Sevilla, Pierre Demeulemeester, Valérie Perhirin, Shuvadeep Dutta, Yashowardhan Sowale, Dinand Tinholt, Emmanuel Poulet-Mathis, Thomas Svahn, Karl Bjurstrom, Monish Suri, Gunnar Menzel, Arne Roßmann, Marek Sowa, Paul Gittins, Luca Bertozzi, Conor MCGovern, Rajesh Iyer, Susanna Östberg, Myriam Chave, Antara Nandy, Mollie Mellows, Raymonde Weyzen, Punam Chavan, Gareth Entwistle, Arghadip Acharya, Aparajita Paul, and Vibha Palekar for their contribution to the report.

## About the Capgemini Research Institute

The Capgemini Research Institute is Capgemini's in-house think tank on all things digital. The Institute publishes research on the impact of digital technologies on large traditional businesses. The team draws on the worldwide network of Capgemini experts and works closely with academic and technology partners. The Institute has dedicated research centers in India, Singapore, the United Kingdom, and the United States. It was recently ranked number one in the world for the quality of its research by independent analysts.

Visit us at [www.capgemini.com/researchinstitute/](http://www.capgemini.com/researchinstitute/)

# For more information, please contact:

## Global contacts

---

**Niraj Parihar**  
[niraj.parihar@capgemini.com](mailto:niraj.parihar@capgemini.com)

---

**Ron Tolido**  
[ron.tolido@capgemini.com](mailto:ron.tolido@capgemini.com)

---

**Anne-Laure Thibaud**  
[annelaure.thibaud@capgemini.com](mailto:annelaure.thibaud@capgemini.com)

---

**Sergey Patsko**  
[sergey.patsko@capgemini.com](mailto:sergey.patsko@capgemini.com)

---

## Local contacts

---

### APAC

**Kirti Jain**  
[kirti.a.jain@capgemini.com](mailto:kirti.a.jain@capgemini.com)

---

### Europe

**Mamadou Sy**  
[mamadou.sy@capgemini.com](mailto:mamadou.sy@capgemini.com)

---

### North America

**Dinand Tinholt**  
[dinand.tinholt@capgemini.com](mailto:dinand.tinholt@capgemini.com)

---

### Financial Services

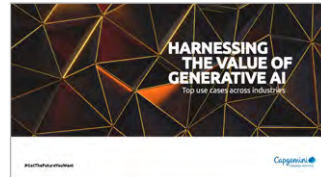
**Ashvin Parmar**  
[ashvin.parmar@capgemini.com](mailto:ashvin.parmar@capgemini.com)

---

# More Capgemini Research Institute Publications



Harnessing the value of generative AI: 2nd edition



Harnessing the value of generative AI: Top use cases across industries



Why consumers love generative AI



Data for net zero: Why data is key to bridging the gap between net zero ambition and action



The data-powered insurer: Unlocking the data premium at speed and scale



The data-powered enterprise: Why organizations must strengthen their data mastery

# Subscribe to latest research from the Capgemini Research Institute



Receive copies of our reports by scanning the QR code or visiting

<https://www.capgemini.com/capgemini-research-institute-subscription/>

Capgemini Research Institute

Fields marked with an \* are required

First Name \*

Last Name \*

Email \*

By submitting this form, I understand that my data will be processed by Capgemini as indicated above and described in the [Terms of use](#).

Submit



A thick, light blue line graphic that starts on the left side of the page, curves downwards into a U-shape, then rises sharply to a peak, and finally curves slightly downwards towards the right. It serves as a background element for the text.

## About Capgemini

Capgemini is a global business and technology transformation partner, helping organizations to accelerate their dual transition to a digital and sustainable world, while creating tangible impact for enterprises and society. It is a responsible and diverse group of 340,000 team members in more than 50 countries. With its strong over 55-year heritage, Capgemini is trusted by its clients to unlock the value of technology to address the entire breadth of their business needs. It delivers end-to-end services and solutions leveraging strengths from strategy and design to engineering, all fueled by its market leading capabilities in AI, cloud and data, combined with its deep industry expertise and partner ecosystem. The Group reported 2023 global revenues of €22.5 billion.

Get the Future You Want | [www.capgemini.com](https://www.capgemini.com)