The data-powered enterprise

Why organizations must strengthen their data mastery

Introduction

The already-rich seam of global data shows no sign of letting up – volumes of digital data are expected to reach 40 ZB this year, up from 1.8 ZB in 2011.[1] To put this into some sort of perspective – this is 40 trillion gigabytes, equivalent not just to every grain of sand on all the world's beaches, but 57 times that amount. However, the sheer volume of data in our hands today in no way translates into equivalent levels of actionable and credible insights.

This is a problem, as data has never been more important. The uncertainty unleashed by the ongoing COVID-19 crisis has heightened the need for forward-looking, real-time insights and scenario modelling, such as pandemic modeling. Actionable data is the most critical factor in digital transformation. But aside from a few high-performers – what we call the Data Masters – our research shows that data capability is patchy among many organizations today.

So, what is it that separates those leading firms that do have a sophisticated capability? What are the unique attributes of their business and technical teams and how can others emulate them to become data-powered enterprises? To answer these questions, we surveyed 500 technology executives and 500 business executives from a range of industries, and conducted detailed interviews with more than 15 executives. We also looked at the relationship between data mastery and the financial performance of organizations.

Based on our research, we look at the following key areas in this report:

1. Where are organizations today in their data journey?
2. What gaps exist between data executives' perceptions and the expectations of business executives on the data mastery of their organizations?
3. Who are the data masters?
4. What are the benefits they achieve?
5. How can we build a data-powered enterprise and what can we learn from the masters in this field?

Executive summary – key takeaways

Organizations have made headway in data-driven decision making and actioning, but the journey is far from over:

- Today, half of organizations say their decision making is completely data-driven, a significant increase on those that said they promoted data-driven decision making in 2018.
- Decision-making approach in organizations, however, is majorly reactive – looking at what happened in the past or why it happened.
- Further, fewer than half (43%) are able to monetize their data and insights through products and services.
- Only a minority (39%) are able to turn data-driven insights into a sustained competitive advantage.

Business executives do not trust their data

- Major gaps exist between the trust that business executives have in the data they receive and how technical executives perceive trust levels. Only 20% of business executives trust the data, but 62% of technical executives believe their business users do trust it.
- Poor data quality is a major contributor to this lack of trust: just 27% of business executives say they are happy with data quality (but, again, there is a mismatch – 54% of technical executives think that their business users are happy with data quality).
- Trusted data is critical for organizational agility and an organization's ability to monetize data.
- While 56% of technical executives believe their data/analytics strategy is aligned with the business strategy, only 38% of business users share that view.
- Similar gaps between technology executives and business executives also exist in areas such as data collection, visibility, and data access policies.

Data masters enjoy significant financial benefits compared to the rest of the sample

- When we assess organizations against critical dimensions of data mastery, only about one in six organizations can be categorized as "data-powered" (we call these the "data masters").

- These data masters significantly outperform the rest of the cohorts on financial parameters such as revenue generation and profitability. Looking at average performance for FY 2019–20, data masters realized a significant performance advantage compared to the rest:
  - 70% higher Revenue per Employee
  - 245% higher Fixed Asset Turnover
  - 22% higher profitability.
- Data masters enjoy a performance advantage of between 30% and 90% in various metrics across customer engagement, top-line benefits, operational efficiency, and cost savings.

How can an organization become a data master?

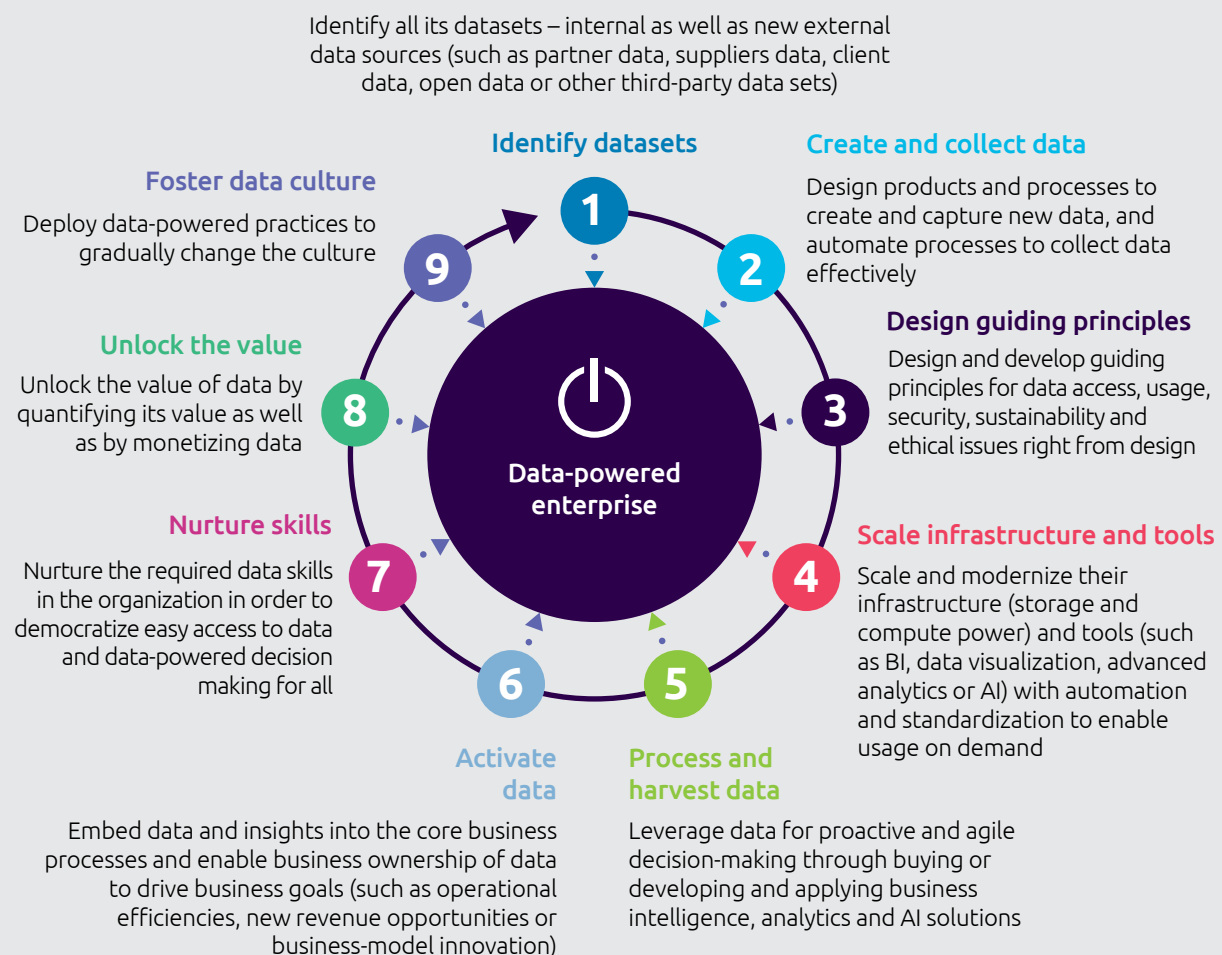Organizations should focus on three priorities:

- **Data activation:** this refers to the use of data in end-to-end business processes to secure business outcomes i.e., transforming data and insights into actions. To achieve this, organizations need to:
  - Align their data and analytics strategy with the business strategy
  - Build business users' trust in data
  - Establish an AI and analytics CoE to assist business teams
  - Foster a data-powered culture by strengthening "data citizenship."
- **Data enablers:** these are the key capabilities necessary for building data-powered enterprises and include infrastructure, governance, and operations. Organizations need to:
  - Strengthen data collection processes and improve data quality
  - Invest in data landscape modernization to get agility in data activation
  - Operationalize data and analytics through DataOps and MLOps
  - Adapt their data governance as data mastery evolves.
- **Data advantage:** this is about creating a competitive advantage, particularly by leveraging external data, including data from hyperscalers. To realize this, organizations need to be able to leverage external data to enhance their insights.

Defining the "data-powered enterprise"

"**Data**" is the digital representation of an organization's past and present, encompassing its processes and interactions with customers, ecosystem, and market.[i]

We define a **data-powered enterprise** as an organization that can create, process and leverage data proactively to fulfill its corporate purpose, achieve its business objectives, and drive innovation.

Figure 1 A data-powered enterprise is able to:

Identify all its datasets – internal as well as new external data sources (such as partner data, suppliers data, client data, open data or other third-party data sets)

Identify datasets

Create and collect data

Design products and processes to create and capture new data, and automate processes to collect data effectively

Foster data culture

Deploy data-powered practices to gradually change the culture

Design guiding principles

Design and develop guiding principles for data access, usage, security, sustainability and ethical issues right from design

Unlock the value

Unlock the value of data by quantifying its value as well as by monetizing data

Data-powered enterprise

Scale infrastructure and tools

Scale and modernize their infrastructure (storage and compute power) and tools (such as BI, data visualization, advanced analytics or AI) with automation and standardization to enable usage on demand

Nurture skills

Nurture the required data skills in the organization in order to democratize easy access to data and data-powered decision making for all

Activate data

Embed data and insights into the core business processes and enable business ownership of data to drive business goals (such as operational efficiencies, new revenue opportunities or business-model innovation)

Process and harvest data

Leverage data for proactive and agile decision-making through buying or developing and applying business intelligence, analytics and AI solutions

Source: Capgemini Research Institute analysis.

[i]**Data & AI platforms** are storage and compute capacities in order to ingest, respond to, govern, transform <u>and use</u> data sets. These platforms help organizations get insights on their business, monitor and report on past and current performance, predict and prescript short-/medium-/long-term actions.

Where do organizations stand today on the journey to data mastery?

Organizations are making headway on data-driven decision making and actioning; yet the approach to using data is still reactive

Given the rapid pace of change in technology, our hyper-competitive business environment, and increasing customer expectations, organizations today need to adapt and reinvent quickly to deal with and even thrive on uncertainty. In this unforgiving environment, harnessing and applying data and analytics is becoming a prerequisite for success and innovation. In our 2018 research, "Understanding digital mastery today"[2] we found this capability to be rare – only 38% of organizations based their decisions on data rather than on intuition or observation.

Organizations have since made some progress on this decision-making aspect. In 2020, 50% of organizations put data at the heart of decision making (see Figure 2). Beatrice Sablone, CDO of Arbetsförmedlingen, a Swedish public employment service, says, *"We want to be much more data driven in our decision making. We want to make more accurate predictions and analysis as an input to political actions that would best benefit the public employment sector. Our clients are our people – basically citizens looking for jobs – and we want to make sure that we help them in the best way – recommending education or jobs or anything that get them out of unemployment as fast as possible. These are our business objectives."*

Figure 2   **Organizations are accelerating data-driven decision making**

Decision making in our organization is data-driven

38% — 2018

50% — 2020
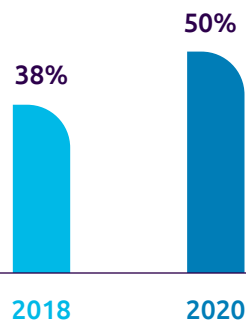
Note: 38% of organizations agreed to the statement "We actively promote data-driven decision making" in 2018; 50% of organizations agreed to the statement, "Decision making in our organization is completely data-driven" in 2020.

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.

We want to be much more data driven in our decision making. We want to make more accurate predictions and analysis as an input to political actions that would best benefit the public employment sector..."

Beatrice Sablone
CDO of Arbetsförmedlingen, a Swedish public employment service

50%

In 2020, 50% of organizations put data at the heart of decision making.

This data-driven capability has never been more important than today, where people need forward-looking insight to steer through the uncharted waters of a pandemic environment. During this crisis, a number of organizations have adopted innovative ways to make the most of external and internal data:

- Dell Technologies has created a real-time dashboard to help its executives make sense of the multitude of data about the COVID-19 crisis and safely guide employees' return to the workplace. John Scimone, senior vice-president and chief security officer at Dell Technologies, says, *"[The dashboard] enables our joint team, with representation from HR, sales, government affairs, security and every part of the company, to look at the same data using the same tool, backed by hard data science and medical intelligence. This ensures our decisions are consistent, data-driven, and informed."* [3]
- A UK-based hospital trust (University Hospitals of Morecambe Bay Trust) has developed an analytical command center that enables its emergency department to see, at a glance, the information they need to make

data-informed decisions. This ranges from the number of ambulances en route to bed capacity.[4]
- Indonesia's largest bank – Bank Mandiri – has built a big-data platform to track transactions and monitor the health of its workforce. During the pandemic, faster access to data and data visualization tools has allowed the bank's management and other stakeholders to be armed with timely information to take informed decisions.[5]

However, while progress has been made on leveraging data for decision making, larger part of the decision-making in organizations still remains reactive that looks at what happened in the past (descriptive) or why it happened (diagnostic). In contrast:

- Only 23% of the time, organizations use predictive approaches (what might happen in the future)
- Only 18% of the time, they use prescriptive approaches (providing recommendations to improve outcomes)
- And just 8% of the time, organizations use an autonomous or self-optimizing approach (systems and processes that help business users make decisions with the objective of achieving a pre-established goal)

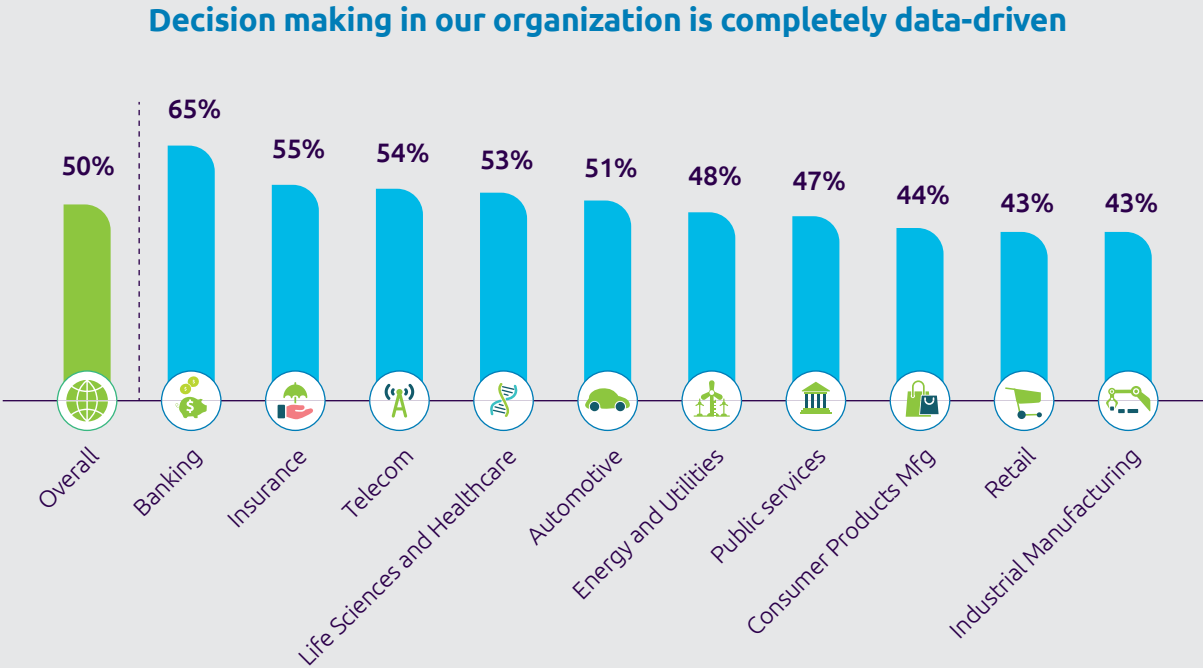**Figure 3**   **Majority of the decision-making in an organization remains reactive**
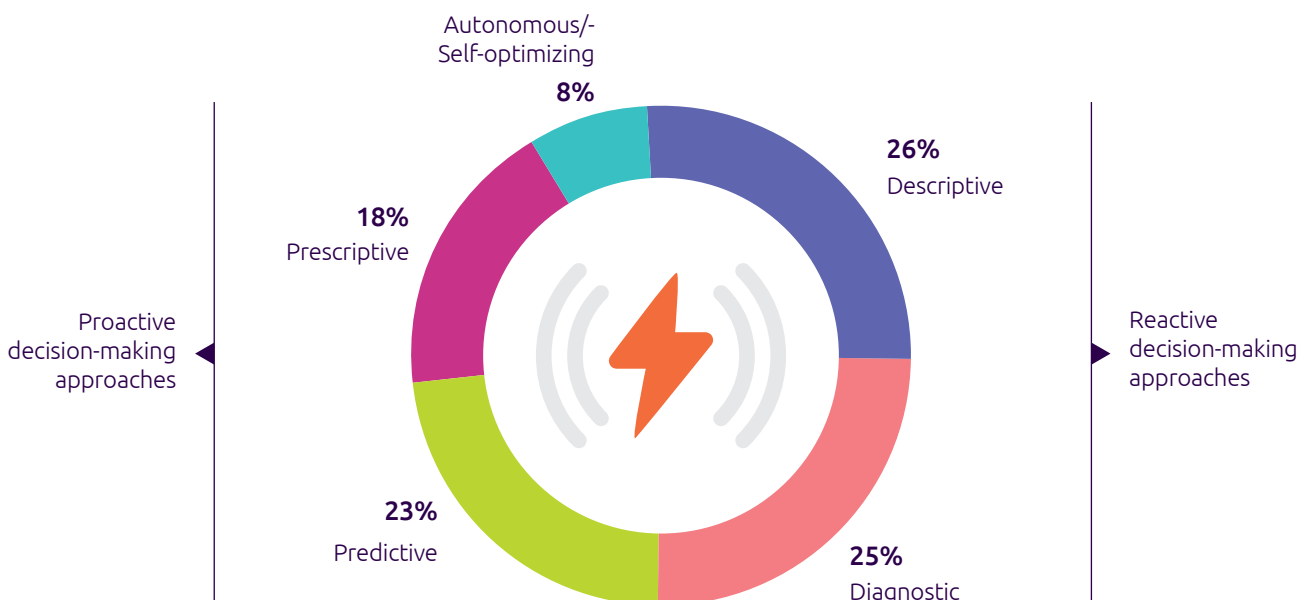
Note: Respondents answered to the question, "Please indicate the proportion of each of the given decision-making approaches in your organization". Reactive decision-making includes "descriptive" and "diagnostic" approaches; while proactive decision-making includes "predictive", prescriptive" and "autonomous/self-optimizing" approaches.

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=262 technology executives who agreed to the statement "decision making in our organization is completely data-driven."

Data-driven decision making: The sector and geographic view

As Figure 4 shows, financial services and telecom sectors lead in using a data-driven approach towards decision making.

Figure 4   **Banking, insurance, and telecom sectors lead in data-driven decision making**
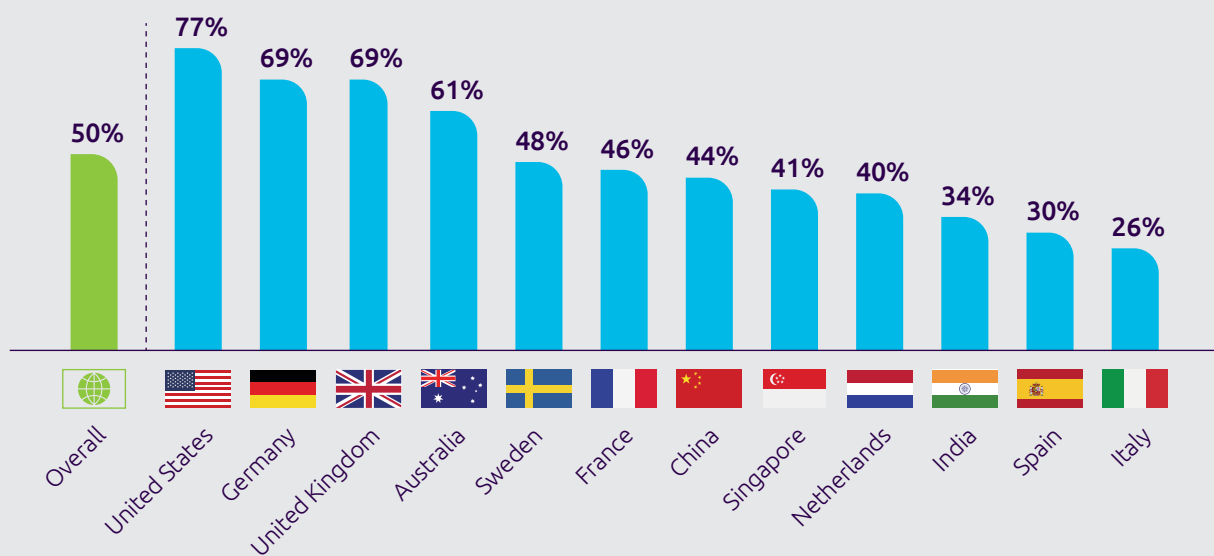
Decision making in our organization is completely data-driven

Overall	Banking	Insurance	Telecom	Life Sciences and Healthcare	Automotive	Energy and Utilities	Public services	Consumer Products Mfg	Retail	Industrial Manufacturing
50%	65%	55%	54%	53%	51%	48%	47%	44%	43%	43%

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=1,004 organizations.

In terms of countries, the US, Germany, and the UK lead the charts, while India, Spain, and Italy trail behind (see Figure 5).
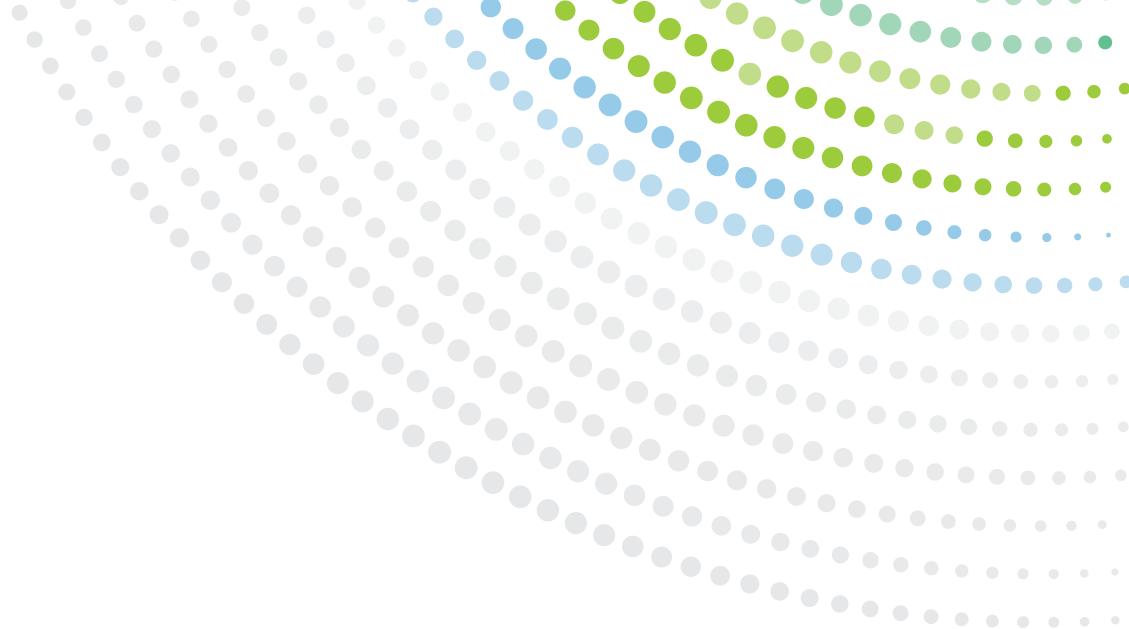
Figure 5   **US, Germany and UK top the charts in data-driven decision making**

Decision making in our organization is completely data-driven

Overall	United States	Germany	United Kingdom	Australia	Sweden	France	China	Singapore	Netherlands	India	Spain	Italy
50%	77%	69%	69%	61%	48%	46%	44%	41%	40%	34%	30%	26%

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=1,004 organizations.

By functions, finance and accounting (67%), risk and compliance (54%), and IT/digital (52%) lead in taking a data-driven approach towards decision making, while functions such as HR (38%) and production/manufacturing (40%) lag behind.

Few have the ability to monetize data or quantify its value

Business and technology executives are almost unanimous in believing that data is one of their most critical strategic assets. Sixty-two percent of the technology and business executives say that data is as an enterprise asset. If data is an asset, it is natural to ask, "how much that data is worth?" and "how can it be monetized?" However, our research reveals that many organizations are not finding answers to those questions. As Figure 6 shows:

- Only 22% of organizations are able to quantify the value of data in their accounting processes. While it may be difficult to do, it is still crucial. The overall approach is about seeing value in terms of how data is used – only when data is applied to a specific use case can organizations determine its value. Christina Ho, former US deputy secretary of commerce, has said, *"I often refer to a simple formula that data plus use equals value."* [6]

The revenue generated by using the data – or the costs that are saved – determine its specific value. For instance, Siemens Mobility understands that scheduled maintenance is cheaper than dealing with an unforeseen breakdown.

The company uses an analytical model to predict maintenance needs, which eliminates costly, unscheduled downtime. That cost of unscheduled downtime becomes the starting point for the price (and value) of the insights service.[7]

- Only 43% are able to monetize data through their products or services (externally or internally). In terms of data monetization, banking and automotive sectors lead:

  – For instance, banks are tying up with retail to monetize their huge data troves. Retail banking customers of Lloyds and Santander can get special offers from a range of retailers. This facility was a results of the two banks joining a digital loyalty scheme run by US-based data advertising firm Cardlytics. The scheme uses spending data to give customers targeted discounts at shops they frequently visit. The banks get a percentage of the fee charged by Cardlytics for running the campaign and Cardlytics gets insights on consumer behavior, which help the retailers to tailor and fund the offers and discounts.[8]

  – In the automotive sector, vehicle data monetization is a big opportunity. Ford CEO, Jim Hackett, has said, *"We have 100 million people in vehicles today that are sitting in Ford blue-oval vehicles. That's the case for monetizing opportunity versus an upstart."* [9]

Figure 6

Only a minority of organizations are able to monetize data and insights

	Overall	Banking	Automotive	Telecom	Insurance	Consumer Products Manufacturing	Public services	Life Sciences and Healthcare	Energy and Utilities	Industrial Manufacturing	Retail
We monetize data assets/insights through our products and services	43%	60%	47%	46%	43%	42%	41%	41%	40%	40%	36%
We quantify the value of data in our accounting systems	22%	28%	21%	30%	19%	14%	26%	18%	25%	16%	19%

■ We monetize data assets/insights through our products and services   ■ We quantify the value of data in our accounting systems

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=1,004 organizations.

To monetize data effectively, organizations need to have a mature approach to designing products and processes that capture new data. In our research, only 44% of organizations say that their products and processes are designed to capture new data. Leading organizations are turning IoT data into new revenue streams. For instance, Otis – one of the largest elevator manufacturers – collects a trove of data on how two billion people move around on any given day, and the company plans to monetize this data by sharing the insights it collects. Chris Smith, vice president of product strategy for Otis, says, *"We can share this data and help others make peoples' experiences more seamless and efficient. Now maybe we could tell coffee shops around a building about what they can expect on a good day, that people will be going there around 11:00 a.m. based on our historical data, so they might want to staff up. That's the kind of knowledge we have that we can start sharing, for a small fee, to help others run their own businesses more efficiently."*[10]

22%

of organizations are able to quantify the value of data in their accounting processes

An interview with Vijay Balakrishnan, Group Chief Data Officer at Michelin

Q:

Our survey reveals that a majority of organizations consider data an important asset, yet they fall behind in quantifying its value or monetizing it. How can organizations quantify the value of the data they have?

Vijay: Let me start with an organization's approach towards any asset – be it people, IP, factories, or machines. Until you get an asset block – you have to invest in it. The next thing you do is that you actually build your business model and strategy around that asset. Thirdly, you maintain the asset. And fourthly, as every asset depreciates, you think about the value coming out of it, its ROI and when you have to replace it.

Now, let me draw a parallel to the data.

- To invest in data assets – organizations need to look at acquiring new data, having the right data governance, cataloguing the data, as well as the need to invest in the required infrastructure of data lakes, including cloud.
- The second thing to do is to chalk out a business plan and strategy along with the data asset. Organizations need to craft the data strategy jointly with various business units. These business units can be at different maturity levels. Some might not have a data strategy; some may have a few BI reporting activities going on; some might be advanced with using AI. So you need to work with them and define a data strategy that is aligned with the overall business strategy. We are targeting the goal of working closely with six Michelin business units over the next year.
- The next one is maintenance. In assets we talk about preventive maintenance, here we think of the feedback loop. So, it is important to have a feedback loop coming from your governance process that identifies where you have bad quality, where you have incomplete data, or where your data needs are not being met, and so on.
- And lastly, you have the ROI. I believe in upfront valuation. We are doing an exercise to get an overall valuation of everything that we are doing in the AI world. Next year, we want to add BI to it, streamline how this valuation is done, and move forward.
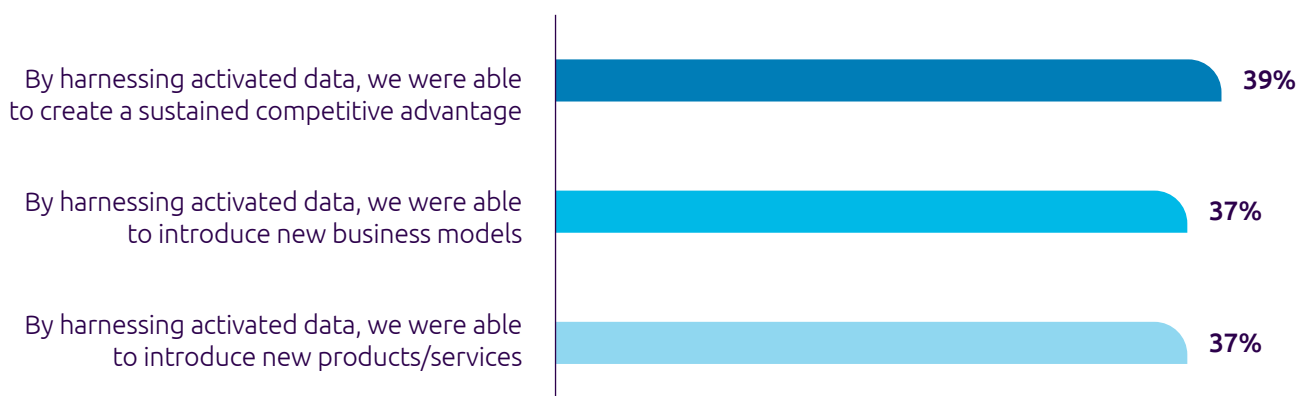
So, I think when any organizations start doing these four aspects, they can really "use" data as an asset.

"Data activation" for innovation is low

By "activated data," we mean driving business value and outcomes through data-driven insights by using algorithms, insights, and intelligent automation for decision making and actioning by the business. It is simply about transforming data and insights into actions within an organization's processes and ways of working that helps them to predict the market dynamics, anticipate trends, identify customer behavior, manage risk, find operational efficiencies, grow, and innovate.

While many organizations have started their data journey, not many have reached the point of truly activating the right data at the right point of action and at the right time. Our research reveals that only a minority of organizations are able to harness the power of "activated data" for introducing new business models, products, or to achieve a sustained competitive advantage (see Figure 7). In terms of sectors, sectors such as banking lead here. For instance, 54% of banking business leaders agree that by activating data in their processes, they were able to create a sustained competitive advantage, whereas only 32% of retail business leaders agreed.

Figure 7   **Most organizations are not able to fully activate data for innovation**

By harnessing activated data, we were able to create a sustained competitive advantage	39%
By harnessing activated data, we were able to introduce new business models	37%
By harnessing activated data, we were able to introduce new products/services	37%

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business executives.
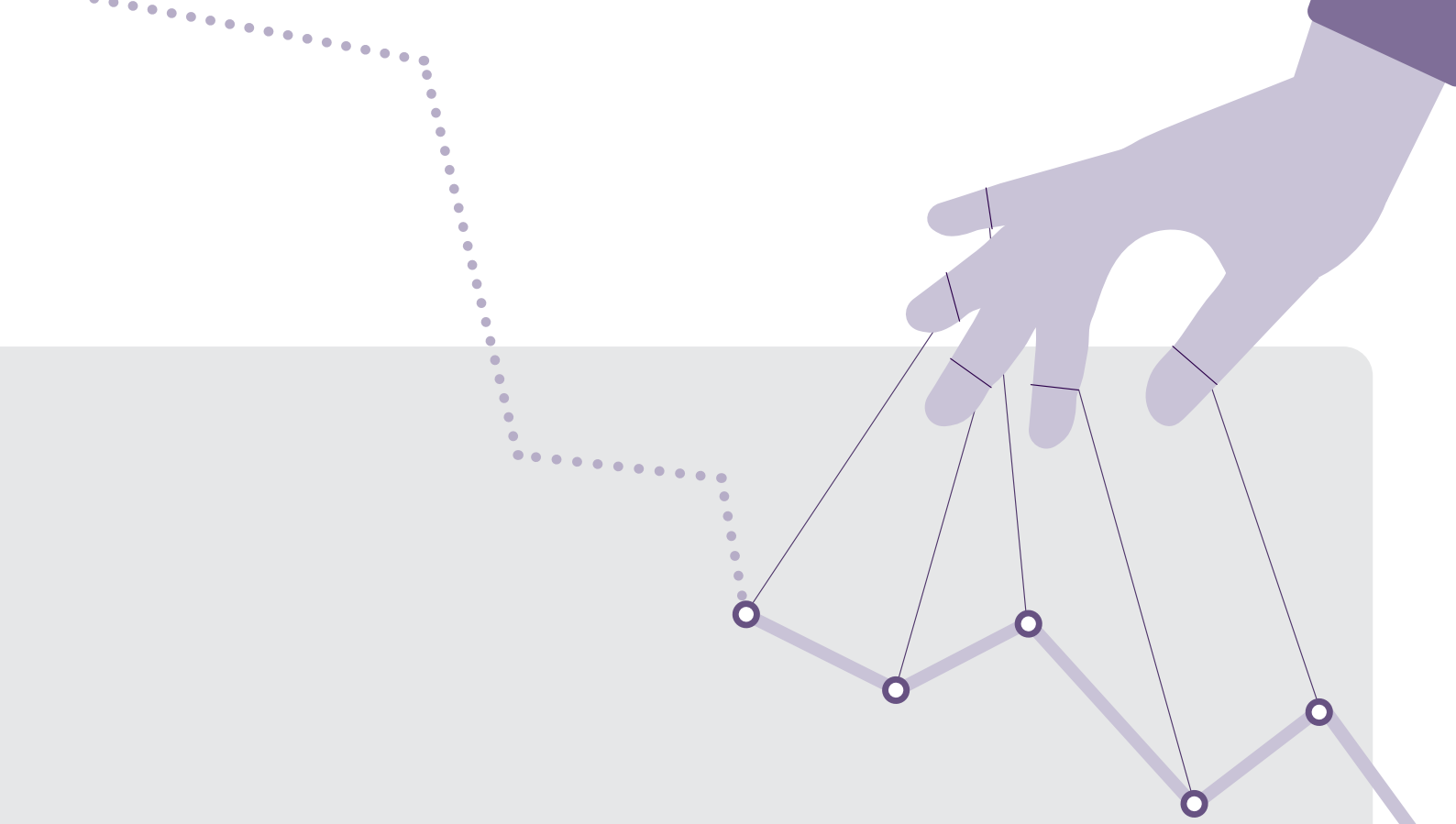
There are a number of mindset and cultural attributes that are critical to activating data:

- Data activation cannot just be the preserve of the technology department. It is critical that it includes the C-suite. The best insights are created in the closest proximity to the business and, to do that, data must be discovered, prepared, analyzed, and visualized right there – and nowhere else.
- Data needs to be viewed as an asset by all employees in the company, rather than something to be dealt with by the IT or a "data" team.
- Openness to testing different proofs of concept to achieve data activation, and having an agile culture of experimentation, are essential. People need to accept and embrace failure.

- Leaders must be willing to have their assumptions and beliefs challenged by data.

Although businesses that are primarily data-powered – such as Uber, Airbnb, etc. – have driven significant success in leveraging the potential of activated data, there are also notable examples in more traditional sectors:

- Michelin has drawn on data-powered customer analysis to provide tires as a "Product-as-a-Service" offering based on distance traveled.[11]
- Paints manufacturer AkzoNobel has a subscription offering for a special paint coat for freighters that reduces fuel consumption. AkzoNobel charges a percentage of the reduced fuel bill.[12]

"Data is business" for platform-based businesses

Data platform businesses such as Uber and Airbnb have been able to build data capabilities that directly link to business strategy, initiatives, processes, and decision support.

Uber: From ride-hailing to a platform

Uber continually analyzes aggregated, anonymous trends to see where their services can run more smoothly. Data science teams from Uber Eats, for instance, analyze how long it takes local restaurants to make a meal, where delivery staff are located, and the effects of weather and time of day to best match orders and routes.[13]

Moreover, based on its data alone, Uber has been able to formulate marketing partnerships with companies such as Starwood Hotels & Resorts, American Airlines, Hilton Hotels, American Express, Capital One, PayPal, and Pepsi.

Data footprint:
still an evolving concept

Given the threat of climate change, the consequences of production, storage, and use of data is a notable sustainability issue. Storage of data consumes large amounts of energy and must be factored into an organization's sustainability strategies. However, our research reveals that organizations lag on this aspect. When we asked business function executives about this issue, only 45% said that sustainability of data (in data production, storage, and access) is a key consideration in defining overall data governance. Countries such as India and Spain trail in this regard, as Figure 8 shows.

A senior executive for a large India-based bank says, *"When we buy something like data storage from vendors, we look at detailed reports on how sustainable they are. This includes areas like their carbon footprint, water usage, policy for disposing of water, and so on. However, we see that most organizations do not care about it. It is also very expensive when you go into that kind of a detail and there are policies from banking regulators that mean you have to store the data for seven years in some cases. Organizations are conservative as a result and store the data for twenty years! I think it is important that data sustainability priorities should trickle down from regulators to organizations."*

Figure 8   **Only a minority consider data sustainability when defining data governance**
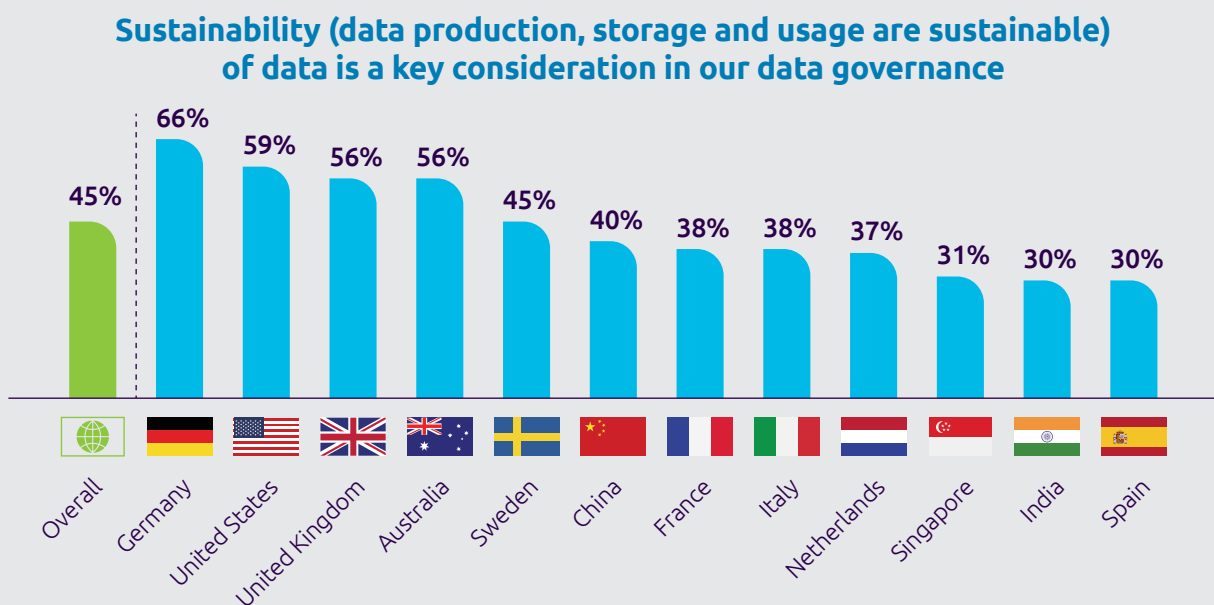
Sustainability (data production, storage and usage are sustainable) of data is a key consideration in our data governance

| Overall	Germany	United States	United Kingdom	Australia	Sweden	China	France	Italy	Netherlands	Singapore	India	Spain
45%	66%	59%	56%	56%	45%	40%	38%	38%	37%	31%	30%	30%

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business executives.

Furthermore, only 50% of the technology executives say that they have established goals to reduce their overall data footprint. As Figure 9 shows, concerted action is also rare.
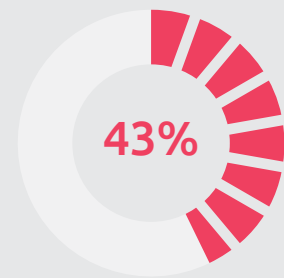
Figure 9 **Only a minority take concerted actions on data sustainability**

39%
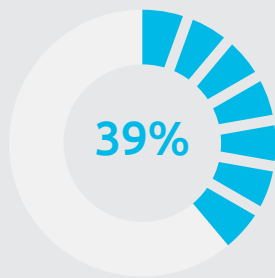
have drafted policies for conserving energy during data storage

42%
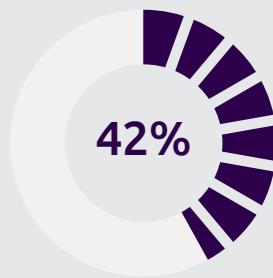
evaluate and audit data vendor partners for their commitment to sustainability

43%

include data energy consumption in the overall carbon footprint

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives.

❝❞

When we buy something like data storage from vendors, we look at detailed reports on how sustainable they are. This includes areas like their carbon footprint, water usage, policy for disposing of water, and so on..."

A senior executive
for a large India-based bank

Significant gaps exist between technology executives' perceptions and the expectations of business executives on data mastery

In our research, we surveyed 500 technology executives and 500 business executives. We found that there is a significant mismatch between the IT and data team's perceptions and the on-the-ground reality for business users.

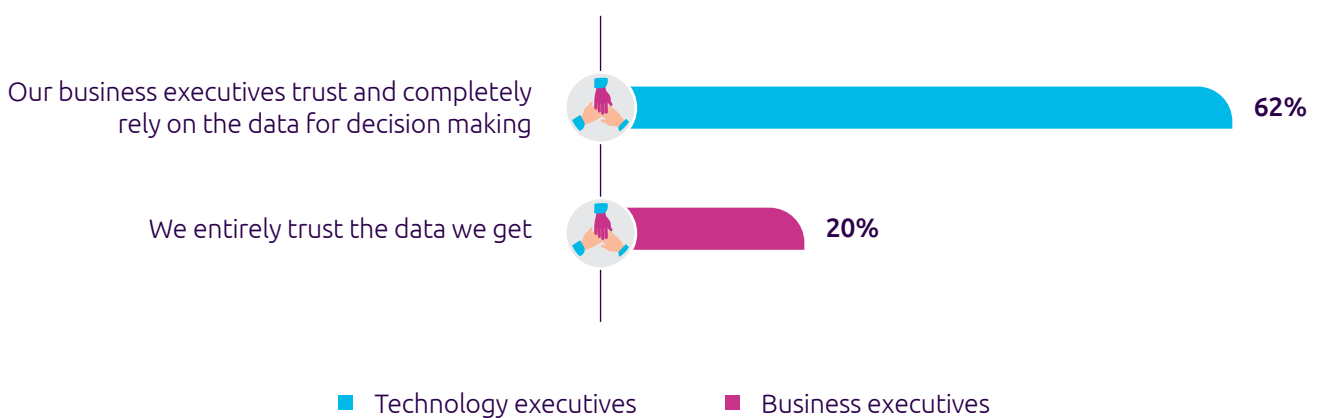## Technology executives lack awareness about the quality and trust concerns of business executives

Business executives still do not trust the data they receive:

To make important strategic decisions based on data, organizations need to trust their data. Trust is the elusive final mile of the data and analytics journey. However, we found a big gap on this aspect. In our survey, 62% of the technology executives say that "our business executives trust and completely rely on the data for decision making." However, only 20% of the business executives say that they have entire trust in the data they receive (see Figure 10).

This lack of trust is particularly pronounced in countries such as India, where only 3% of business executives say that they trust the data they get, followed by the Netherlands (9%) and Italy (13%). In terms of sectors, only 6% of business unit executives in retail and 14% in public services and insurance sectors trust the data they get.

Figure 10 **Business executives do not trust the data they get**

Our business executives trust and completely rely on the data for decision making — **62%**

We entirely trust the data we get — **20%**

■ Technology executives ■ Business executives

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives, N=504 business unit executives.

20%

of business executives completely trust the data they get

Data quality is a big concern for business:

Data quality is superset – and ranges from data sourcing quality issues (garbage in, garbage out), technical and business quality issues, to data veracity issues (accuracy, precision, and trustworthiness of data). Poor data quality is one of the main contributors in this lack of trust. Data quality is not often thought of as the most exciting aspect of data engineering and data science, but that does not take away from its critical importance.

The consequences of poor data quality can be enormous:

- **In a research study published in MIT Sloan Management Review, companies are said to be losing around 15% to 25% of their revenue due to poor data quality**. These costs come as people accommodate bad data by correcting errors, seek confirmation from other sources, and deal with the inevitable mistakes that follow.[14]
- Another consequence of poor data quality is that **knowledge workers waste up to 50% of their time dealing with mundane data-quality issues**.[15]

And, our survey reveals that data quality is a huge issue for business executives today, which technology executives are unaware of. As Figure 11 shows, 54% of technology executives think that their business users are happy with data quality. However, only 27% of business executives say that they are happy with data quality.

| Figure 11 | Data quality is a huge issue for business executives |
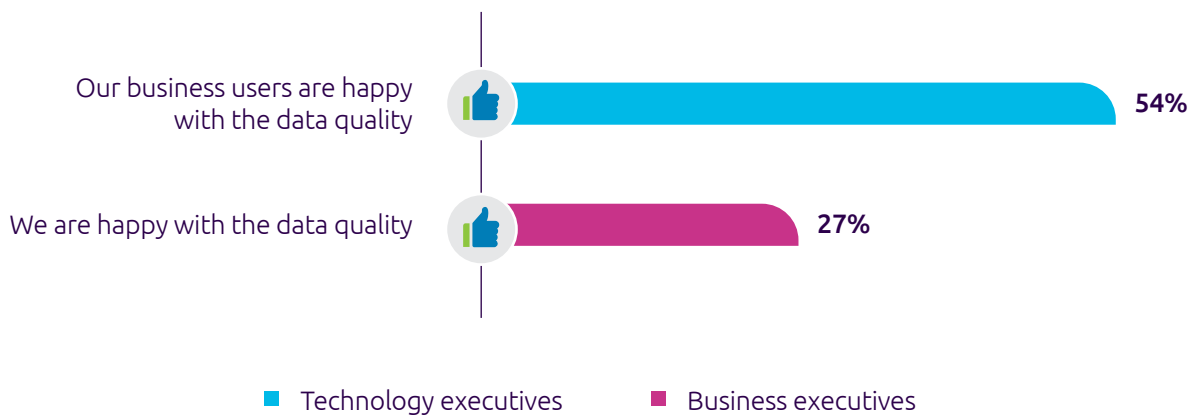
Our business users are happy with the data quality — **54%**

We are happy with the data quality — **27%**

■ Technology executives   ■ Business executives

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives, N=504 business unit executives.

> *Data quality is the top priority for us. We have made progress on the availability of our system for the last three years, and now we are working on data quality, which can be the most important cause of dissatisfaction among users..."*

Yves Caseau,
Group CIO for Michelin Group

Leading organizations such as Michelin have understood this challenge and have started to tackle data quality head-on. Yves Caseau, Group CIO for Michelin Group, told us, *"Data quality is the top priority for us. We have made progress on the availability of our system for the last three years, and now we are working on data quality, which can be the most important cause of dissatisfaction among users. We are working on three things. Firstly, our data model, where our goal is to come up with a federated model, as some of our data quality problems comes from the fact that the data models are not properly understood. Secondly, we are looking at systems which are intuitive for people to access. And the third domain is business process operations transformation – to simplify processes that are complex or inconsistent while working on the same data."*

Data quality is also essential to fully exploit the possibilities of artificial intelligence (AI). Our previous research on AI-powered enterprises highlights that **"improving data quality" ranks as the number-one approach that AI-at-scale leaders use to get more benefits from their AI systems.**[16]

Addressing data quality issues paves way for a clear ownership of data. *"Empowered and accountable data ownership is what organizations need and should be taken seriously from end to end,"* says Kevin Fletcher, CDO & Chief Economist and Tim East, CIO, HM Revenue & Customs. *"This also includes managing the suppliers [of data] in the supply chain. This ownership vests in the partnership between the data owner that sits in the business, the CIO who sits in our [technical] area and the partner that is working to get the effective data delivered. And sometimes, that relationship gets too distant and that is when the empowered and accountable ownership gets diluted."*

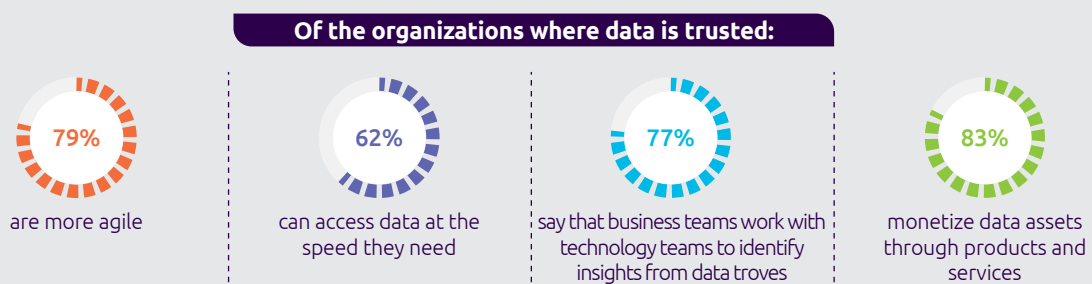# Data trust is key to organizational agility, collaboration and generating value

We analyzed those business executives in our survey who said they completely trust the data they receive. This revealed some notable differences between this cohort and those who do not trust the data they receive.
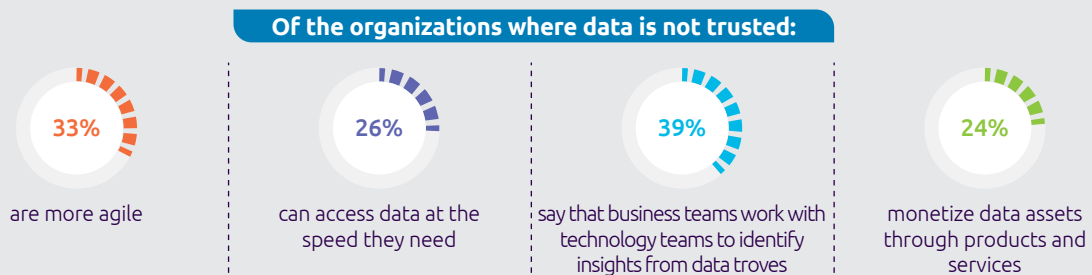
As Figure 12 shows:

- **Organizations where business executives trust data are more agile:** Of the organizations where data is trusted, 79% of those organizations consider themselves more agile, and 62% say that they can access the data at the speed they need. However, out of the organizations where data is not trusted, only 33% see themselves as agile.

- **Organizations where business executives trust data are better able to drive cross-functional collaboration to leverage insights:** Of the organizations where data is trusted, 77% of those organizations boast of better cross-functional collaboration.

- **Organizations where business executives trust data are better able to monetize data:** Of the organizations where data is not trusted, only 24% of those organizations are able to monetize data in comparison to 83% where data is trusted.

Figure 12 Organizations where data is trusted are more agile, collaborative, and are able to unlock value from data

Of the organizations where data is trusted:

79%	62%	77%	83%
are more agile	can access data at the speed they need	say that business teams work with technology teams to identify insights from data troves	monetize data assets through products and services

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=102 business executives who disagree to the statement, "we do not entirely trust the data we get".

Of the organizations where data is not trusted:

33%	26%	39%	24%
are more agile	can access data at the speed they need	say that business teams work with technology teams to identify insights from data troves	monetize data assets through products and services

Source: 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".
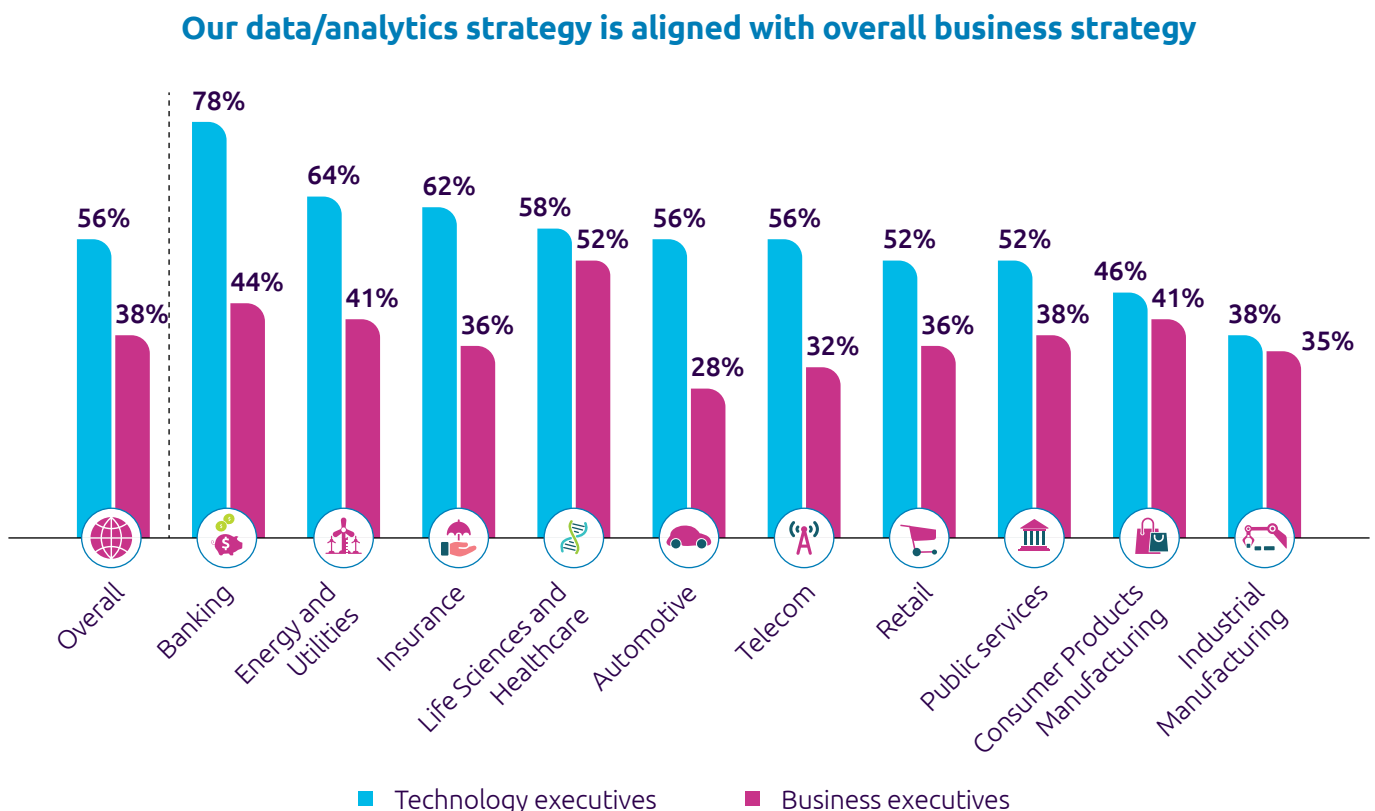
Business and technology teams differ on whether data strategy is aligned with organizational vision

Building an enterprise-wide data strategy which is aligned with organizational vision is crucial for success of any data initiative. Nitin Sethi, global IT director for business transformation & engagement at Visteon Corporation, an automotive supplier, says, *"The key is to build an enterprise-wide data strategy. This is very important to embark on the data journey. Solid data strategy should comprehend vision and road map for an organization's potential to harness data-dependent capabilities and should be strictly outcome-driven."*

However, as Figure 13 shows, 56% of technology executives believe that data/analytics strategy is aligned with the overall business strategy. But, only 38% business executives feel the same. This emphasizes the mismatch in organizations on not only defining the strategy, but also on designing the architecture and deploying the platform and operating model. Successful organizations have moved from "data strategy supports the business strategy" to "data strategy is the business strategy." The mismatch is particularly stark in the banking industry.

Figure 13 | **Business executives feel a gap in alignment of data and business strategy**

Our data/analytics strategy is aligned with overall business strategy

Technology executives | Business executives

	Overall	Banking	Energy and Utilities	Insurance	Life Sciences and Healthcare	Automotive	Telecom	Retail	Public services	Consumer Products Manufacturing	Industrial Manufacturing
Technology executives	56%	78%	64%	62%	58%	56%	56%	52%	52%	46%	38%
Business executives	38%	44%	41%	36%	52%	28%	32%	36%	38%	41%	35%

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives, N=504 business unit executives.

Technology executives are unfamiliar with the concerns that business users have about data visibility, access, and insights
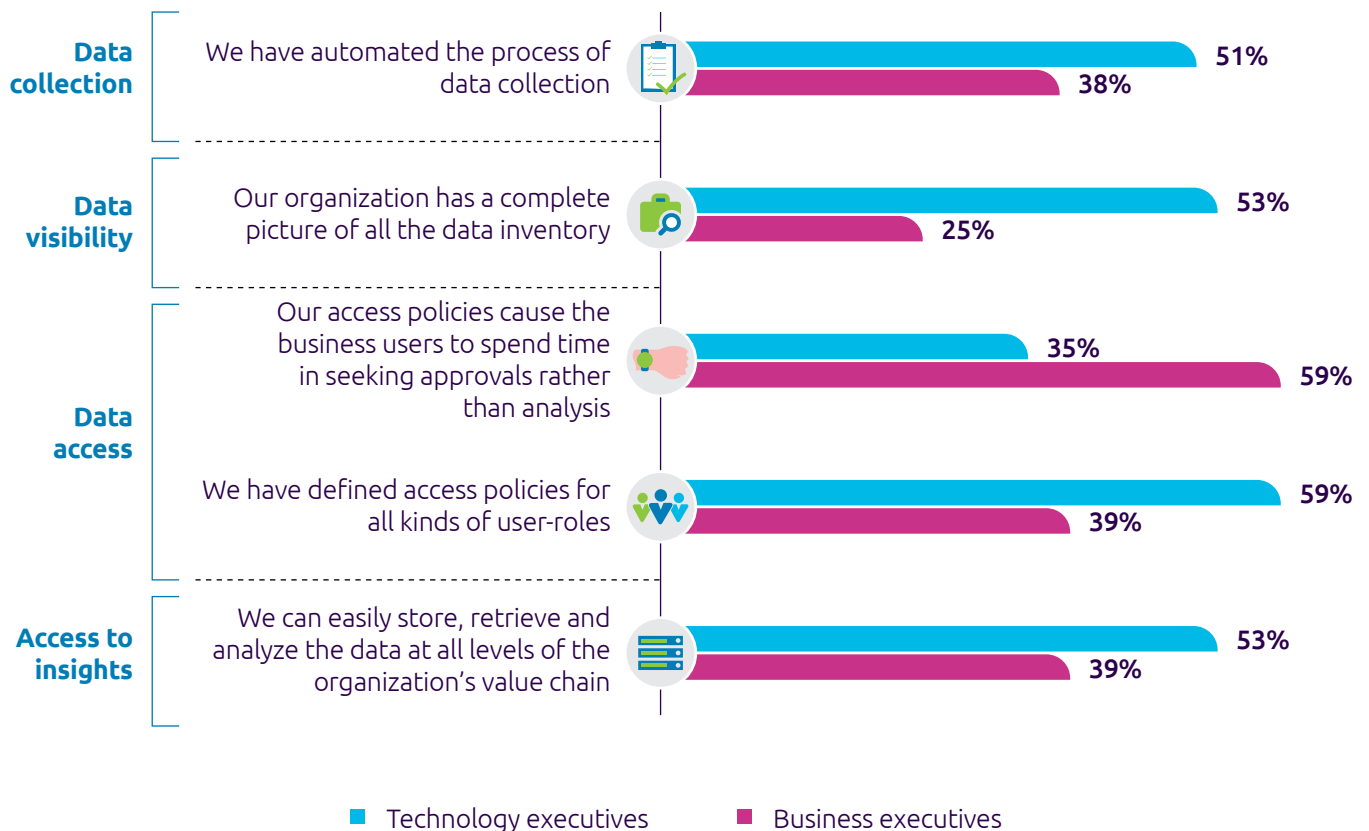
Figure 14 shows the extent of the mismatch:

- Data collection is a key challenge for organizations. Specifically for manufacturing sector, our previous research on smart factories has revealed that platforms for collecting, aggregating, and analyzing data are low in maturity. Only 46% of manufacturers have a horizontal data collection and aggregation platform and only 40%

have an end-to-end integrated platform that reaches from device to analytics.[17] Our current research also shows a gap on this aspect – 51% of technology executives say that they have automated the process of data collection, whereas only 38% of business executives agree.

- 53% of technology executives say that their organization has a complete picture of all the data inventory. However, only 25% of business executives agree.
- Getting access to data at the speed which business needs is critical to effective decision making and to building trust. However, there are significant gaps on clarity of access policies and ease of gaining access to the data required for decision making.
- In terms of retrieving and analyzing the data at all levels of the organization's value chain, there is a significant gap of 15% points between technology and business executives.

Figure 14 **Data collection, visibility, and access are pain areas for business executives**

		Technology executives	Business executives
Data collection	We have automated the process of data collection	51%	38%
Data visibility	Our organization has a complete picture of all the data inventory	53%	25%
Data access	Our access policies cause the business users to spend time in seeking approvals rather than analysis	35%	59%
Data access	We have defined access policies for all kinds of user-roles	59%	39%
Access to insights	We can easily store, retrieve and analyze the data at all levels of the organization's value chain	53%	39%

■ Technology executives   ■ Business executives

> *The key is to build an enterprise-wide data strategy. This is very important to embark on the data journey. Solid data strategy should comprehend vision and road map for an organization's potential to harness data-dependent capabilities and should be strictly outcome-driven."*

Nitin Sethi,
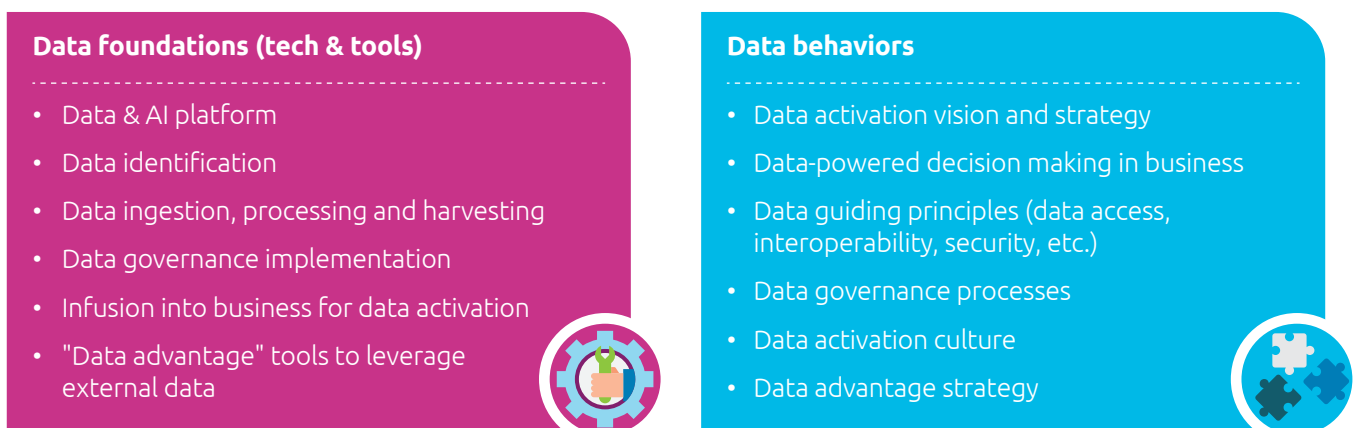global IT director for business transformation & engagement at Visteon Corporation, an automotive supplier

Who are the data masters?

Understanding our data mastery framework

In order to understand where organizations are on the journey to being data powered, we analyzed the organizations that participated in our research on a number of critical elements.

As Figure 15 shows, 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 relates to people, processes, skills, and culture. Taken together, they drive data mastery.

Figure 15 **Elements of data mastery**

Data foundations (tech & tools)

- Data & AI platform
- Data identification
- Data ingestion, processing and harvesting
- Data governance implementation
- Infusion into business for data activation
- "Data advantage" tools to leverage external data

Data behaviors

- Data activation vision and strategy
- Data-powered decision making in business
- Data guiding principles (data access, interoperability, security, etc.)
- Data governance processes
- Data activation culture
- Data advantage strategy

Source: Capgemini Research Institute analysis.

*More details on sub-parameters included for both the dimensions are mentioned in Appendix 2.

Based on these elements of data mastery, we identified four different cohorts. Of these:

16%

of the organizations we surveyed fall into the mature or "data masters" category, where they lead in both data behaviors and facilitators
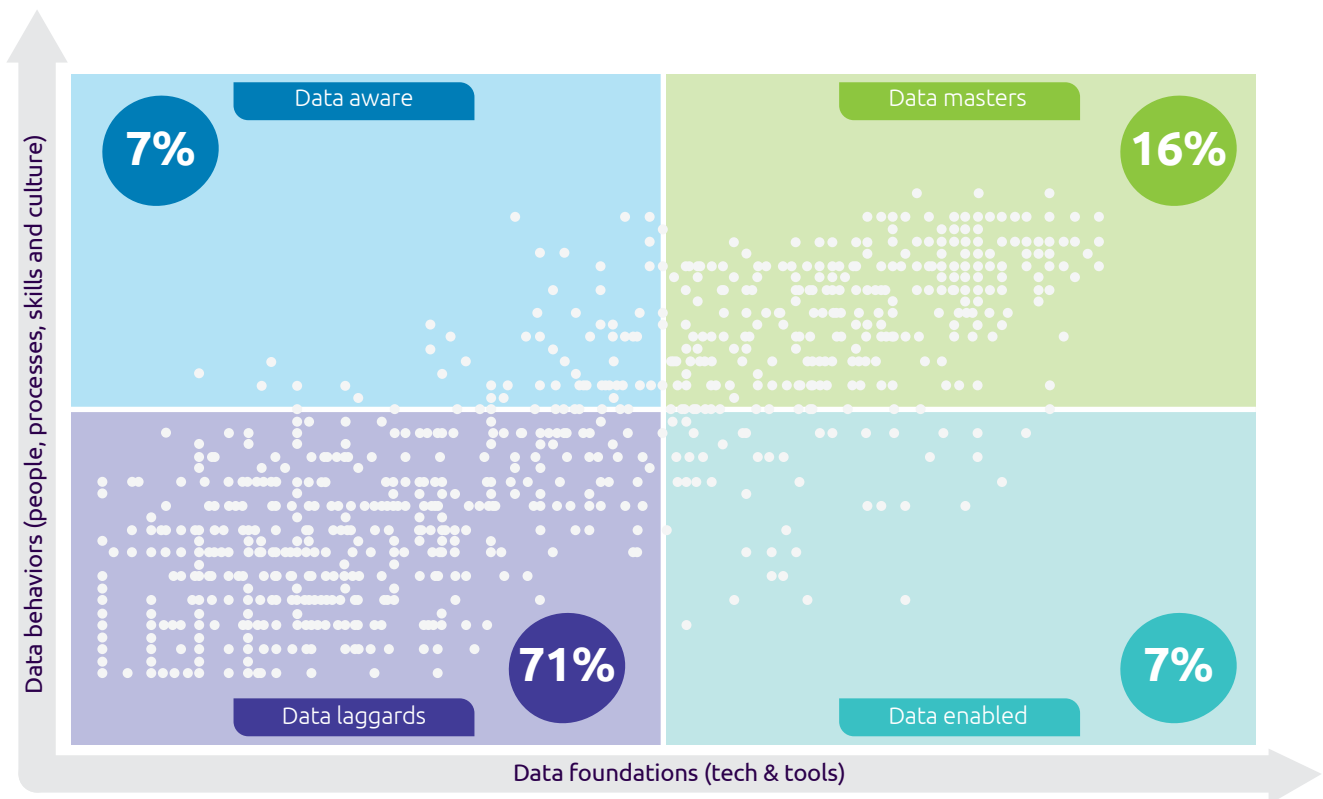
71%

fall into the "data laggards" category

14%

do well in one of the dimensions but not the other.

Figure 16 **Only about one in six organizations can be categorized as a data master**

Source: Capgemini Research Institute analysis, Data-powered enterprises survey, August 2020, N=500 technology executives, N=504 business unit executives. Percentages are rounded off to the nearest integer and do not sum to 100.

Data masters vs the other cohorts: Key characteristics

Analyzing the data laggards, data enabled, data aware, and data masters, we highlight key characteristics of these cohorts in terms of people, process, technology, data monetization and other features (see figure 17).

Figure 17 Key characteristics of the four cohorts

Cohort	Data laggards	Data enabled	Data aware	Data masters
Features	Basic reporting capabilities and minimal visualization	Data and the infra for reporting and visualization exists but no clear capabilities on effectively using this for business	Business is keen on leveraging data and insights but are handicapped due to issues of data quality or skills or governance	Advanced reporting and analytical capabilities exist in each business area
Data discovery	Challenging for business users and new staff to understand the data	Data discovery is complex due to restricting policies around access and usage	Data discovery is complex due to lack of strong master data management	Well maintained master data and data catalogues allowing easy data discovery
Relationship between technical and business users	Business depends on technical teams for all reporting needs	Data is available for analysis, but business is not able to leverage this for insights	Business users are aware of the kind of analyses they need but they are not easy to obtain	Business and technical teams collaborate to understand how they can reach business objectives
People	Technical teams handle business requirements	Despite the availability of the required tools and platforms, the business is dependent on the technical teams	Each BU has dedicated resource(s) that can run simple queries and extract reports	Business users are empowered to build their own reports and dashboards

Cohort	Data laggards	Data enabled	Data aware	Data masters
Training	Minimal focus on upskilling technical or business users	Training on data and analytics is available to technical users but minimal training exists for business users	Training programs exist for selected business users	Continuous focus on upskilling majority of the business users
Processes	Minimal guidelines on data access, privacy and none on ethics	Guidelines on data access, use, and privacy exist, but limit the ability of business to explore data	Guidelines on data access, use, privacy, security exist but no standards defined for ethics	Robust framework covering all aspects of privacy, security, ethics apart from access and use
Technology for data activation	Siloed data sources	Integrated data platforms exist alongside legacy platforms, which makes data discovery and exploration confusing	Integrated data platforms exist but querying and reporting takes time	Unified data platforms leveraging cloud and self-service tools
Data advantage (Use of external data sources)	No thinking on how to leverage external data	Data platforms make the external data sources available for analysis, but cross-analysis is not facilitated	Business looks at external data but integrating with other types of data is still complex	Actively work with external players to collect various sorts of data and leverage this for insights
Data monetization	No opportunities to monetize data	Monetization is not seen as an opportunity	Business is keen on monetizing however it is not supported	Data is an integral part of decision making and can even monetize data

Source: Capgemini Research Institute analysis.

The data mastery advantage

Data masters drive more profits and revenue compared to the other cohorts

We analyzed the 2019–20 financials, by industry, for each of the 700+ unique organizations in our research to understand whether data mastery translates to a better financial performance. As Figure 20 shows, we found that the data masters, on average, outperform the rest of the cohorts on multiple financial metrics. When we looked at the metrics for FY 2018–19, the data masters also enjoyed a superior performance.

22%

Data masters are 22% more profitable than the average organization

Revenue generation:

Revenue/employee: the figure in the cohort indicates the difference in performance of that particular cohort from the average value for all the organizations. For instance, the data masters have seen a 70% higher "revenue per employee" compared to the average, while the data laggards have realized a 16% lower "revenue per employee".

Fixed Asset Turnover: Data masters realized 245% higher "Fixed Asset Turnover" compared to the average.
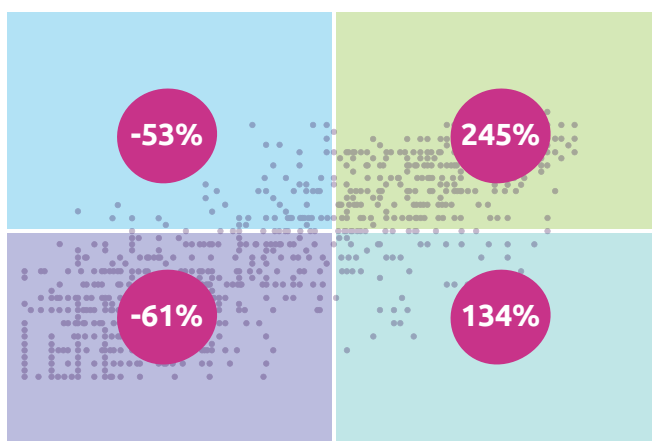
Profitability:

Profitability is computed from both EBIT Margin and Net Profit Margin. Here, we found that the "data-aware" cohort lags the beginners in the profitability by 67% compared to the average. However, the data masters are 22% more profitable than the average.
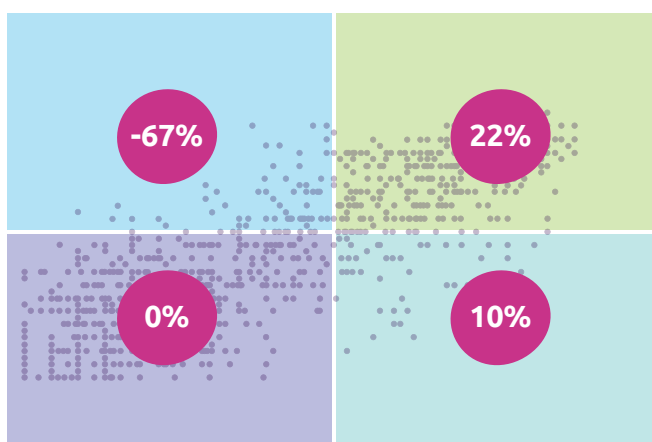
Source: Capgemini Research Institute financial analysis of 739 unique companies (N=111 data masters, 41 data aware, 53 data enabled and 534 data laggards) for FY 2019–20.

Data masters' superior financial performance a result of a performance advantage

The strong financial performance of data masters reflects their performance advantage compared to the others in the sample.

Customer engagement is higher in data-powered enterprises than in others

In the following sections, we look at how data masters performed compared to all others (data enabled, data aware, and data laggards). We found that data masters realized 30%–90% higher performance compared to others in various business metrics.

For example, as Figure 19 shows, **data masters achieved a 22% reduction in customer churn, which is 87% more than what others realized (12%).**

Figure 19   Customer engagement benefits realized by being data powered

Extent of customer engagement related benefits seen
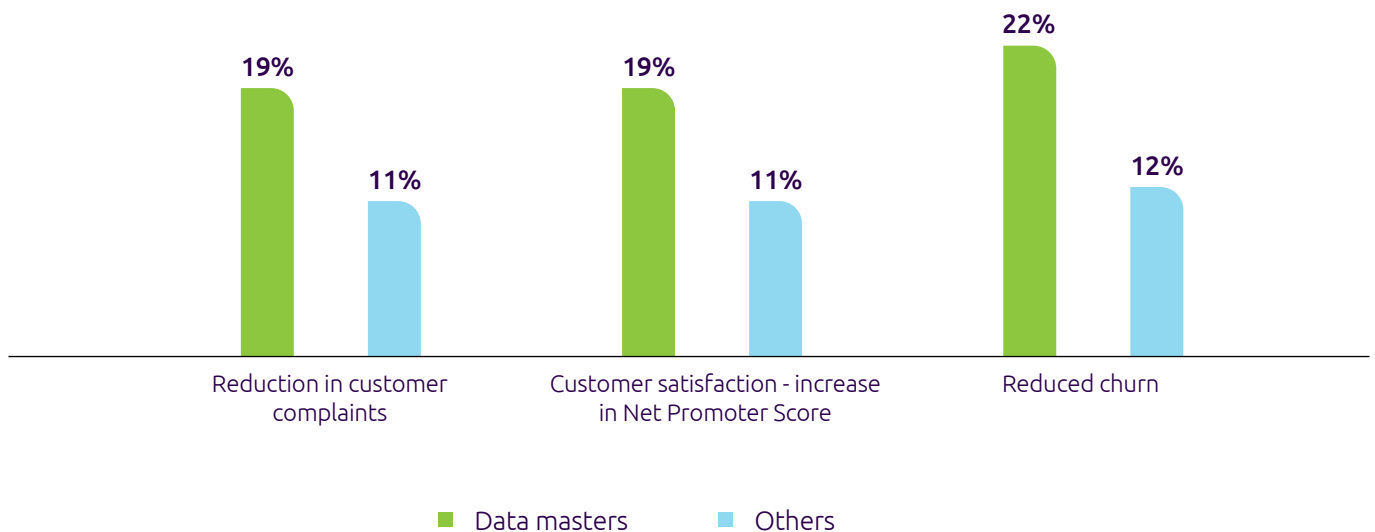
- **Reduction in customer complaints:** Data masters 19%, Others 11%
- **Customer satisfaction - increase in Net Promoter Score:** Data masters 19%, Others 11%
- **Reduced churn:** Data masters 22%, Others 12%

■ Data masters   ■ Others

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business unit executives, N=72 business data masters; N=432 business others.

Organizations can use data to enrich their customer engagement in a number of ways:
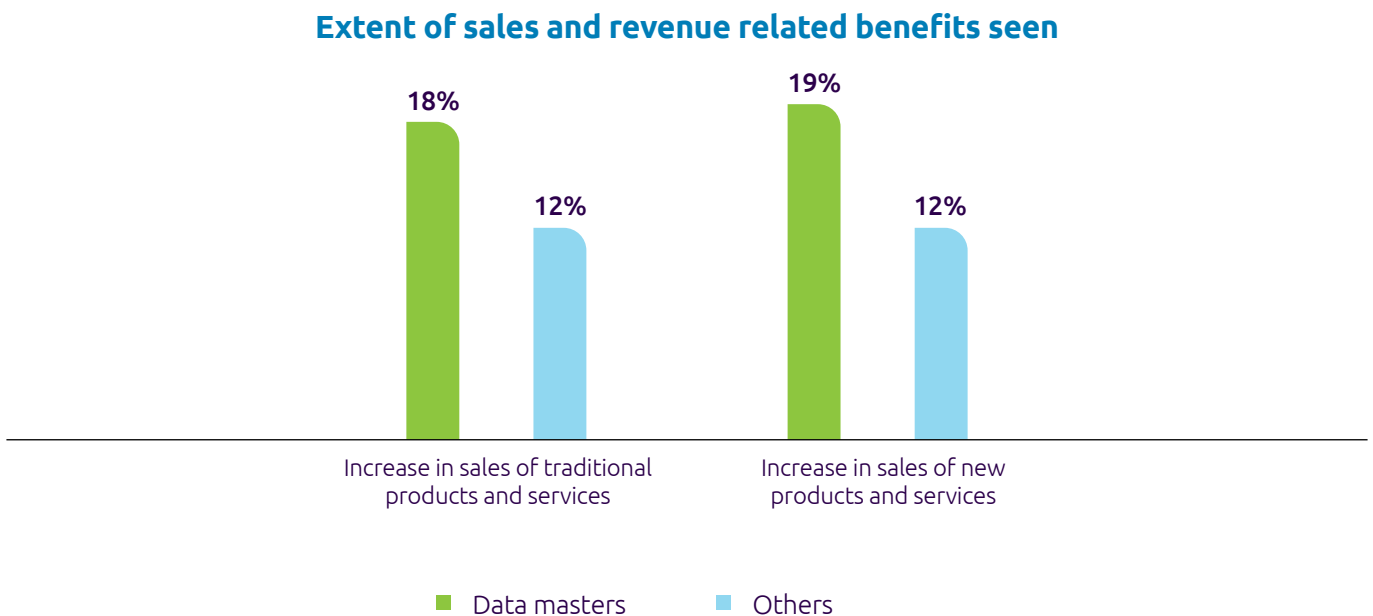
- Training bots on the most frequently asked customer service questions to reduce waiting times for customers. Royal Bank of Scotland also uses data-based patterns to predict what products a customer may want and what issues may arise. Real-time data analytics help the company track complaints and predict what questions or complaints customers may have.[18]

- Analyzing consumer preferences to come up with better products or new features.
- Personalizing customer interactions and experience by storing the history of interactions and the customer's pain points. Highlighting the need for data, Eric Sprunk, COO of Nike, says, *"As demand for our product grows, we must be insights driven, data optimized and hyper-focused on consumer behavior. This is how we serve consumers more personally at scale."* [19]

Data-powered enterprises lead in top-line benefits

Insights from data can also help organizations in sales of their products and services. Unilever, for example, identified emerging consumer interest in "ruby chocolate" (i.e., pink chocolate as opposed to white or dark) through its social business analytics platform. This insight allowed Unilever to launch the first variant in this space ahead of its competitors.[20] Similarly, Conagra Brands' AI platform helped the company identify a new segment, "flexitarians" (flexible vegetarians), who were also interested in compostable packaging, by analyzing the data from social media sites and market-research companies. This led Conagra to introduce grain-free varieties of its Healthy Choice Power Bowl in compostable containers.[21] As Figure 20 shows, our research showed that data masters realize a 19% increase in sales of new products and services compared to 12% for the rest.

Figure 20   **Top-line benefits realized by being data-powered**

Extent of sales and revenue related benefits seen

- 18% — Increase in sales of traditional products and services (Data masters)
- 12% — Increase in sales of traditional products and services (Others)
- 19% — Increase in sales of new products and services (Data masters)
- 12% — Increase in sales of new products and services (Others)

■ Data masters   ■ Others

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business unit executives, N=72 business data masters; N=432 business others.
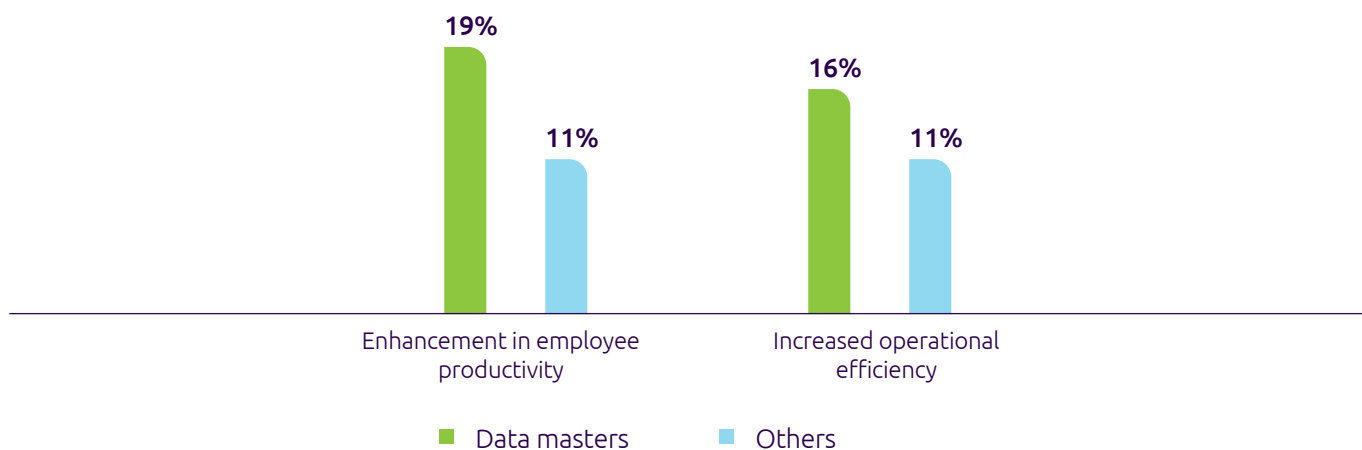
Data-powered enterprises gain greater operational efficiencies

Leveraging the insights from data, organizations can make significant gains in operational efficiency and employee productivity (see Figure 21). For instance, Volvo's "Early Warning System" is powered by large datasets. Every week, the system analyzes over one million events that occur during machine operations, such as temperature increases or abnormal pressure readings. This allows the organization to assess their impact on breakdown and failure rates.[22]

Figure 21 **Operational benefits realized by being data-powered**

Extent of operational benefits seen

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business unit executives, N=72 business data masters; N=432 business others.

Data-powered enterprises drive greater cost savings

Our research shows that, on average, the data masters used data-powered insights to realize a 15% gain in operational cost reduction resulting from process improvements,

and a 17% savings gain from fraud detection (see Figure 22). To illustrate, Danske Bank recently modernized its fraud detection process and reduced the amount of fraud detected by 1,200 claims a day. By revamping this process, the bank was able to reduce its false positives by 60% – and is expected to reach 80% reduction as the new machine learning model continues to learn.[23]

Figure 22 **Savings realized by being data powered**

Extent of cost-related benefits seen
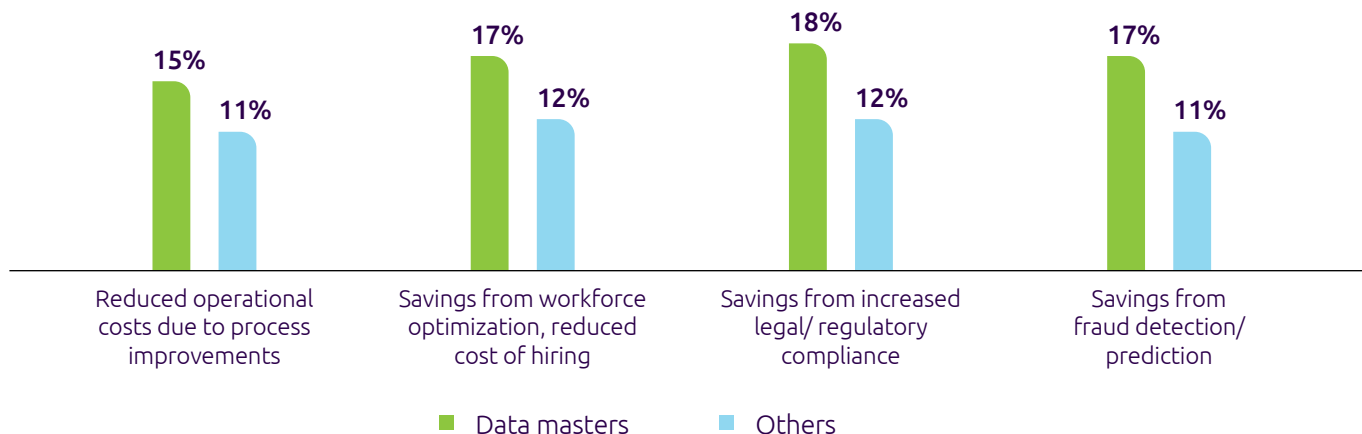
Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business unit executives, N=72 business data masters; N=432 business others.

A list of use cases for each of the business objectives above is given in Appendix 1.
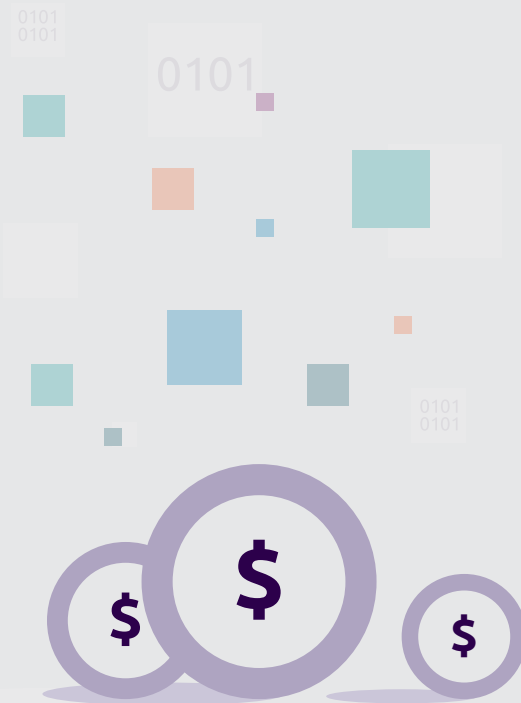
Data for good: a social enterprise for east African farmers

Organizations are using data-powered innovation to tackle environmental and social issues. Agrics is a social enterprise operating out of East Africa that offers agricultural products and services on credit to local famers. Together with Agrics, Capgemini developed the intelligent data platform – FARM (Financial & Agricultural Recommendation Models) – to address the challenges that the farming community faced both in terms of lack of resources and the lack of connectivity. The platform uses AI to determine farming patterns through big data, generating insights from the information to make recommendations. It is available at scale by connecting it with cell phones. The platform collects data from a number of public and private resources and hosts it in the cloud. It can fulfill a number of functions:

• Advising farmers on their expected crop yield using pictures of the farmer's soil and machine learning models to classify soil quality. Combining this with weather data, soil maps, satellite data and crop models, the platform offers tailor-made advice to farmers
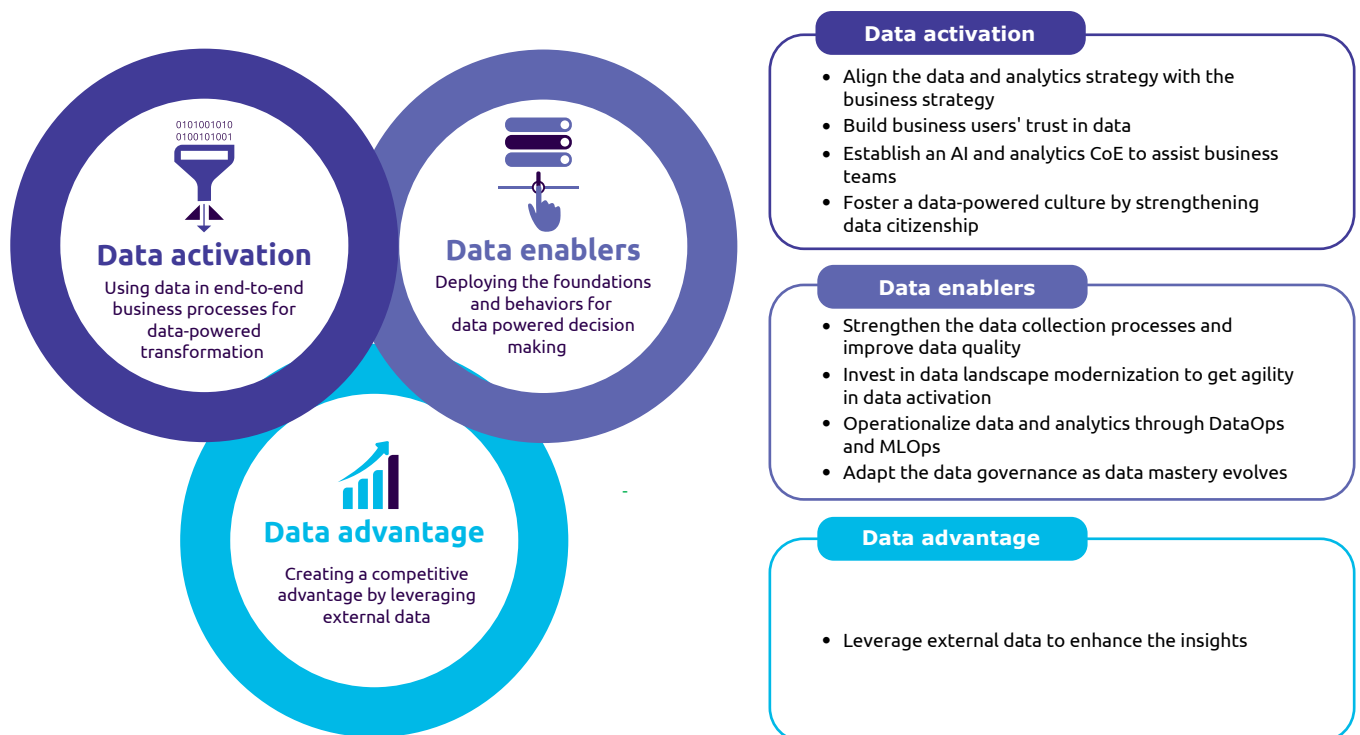• Advising Agrics on harvest expectations and repayment risk.

By bringing together different players in the value chain, the platform enables local communities of farmers to optimize their agricultural processes.[24]

Building a data-powered enterprise: Learning from the data masters

Drawing on our research, as well as our experience working with organizations in this space, we believe three capabilities are critical for building a data-powered enterprise (see Figure 23):

1. Data activation (the use of data in end-to-end business processes to secure business outcomes)
2. Data enablers (deploying the foundations and behaviors for data-powered decision making and include infrastructure, governance, and operations)
3. Data advantage (creating a competitive advantage, in particular by leveraging external data, including data from hyperscalers).

Figure 23   **Three areas of focus to become a data-powered enterprise**

Data activation
Using data in end-to-end business processes for data-powered transformation

Data enablers
Deploying the foundations and behaviors for data powered decision making

Data advantage
Creating a competitive advantage by leveraging external data

Data activation
- Align the data and analytics strategy with the business strategy
- Build business users' trust in data
- Establish an AI and analytics CoE to assist business teams
- Foster a data-powered culture by strengthening data citizenship

Data enablers
- Strengthen the data collection processes and improve data quality
- Invest in data landscape modernization to get agility in data activation
- Operationalize data and analytics through DataOps and MLOps
- Adapt the data governance as data mastery evolves

Data advantage
- Leverage external data to enhance the insights

Source: Capgemini Research Institute analysis.

Data activation

1 Align the data and analytics strategy with the business strategy

Figure 13 has highlighted the misalignment between the technical executives and business executives on the fit of data/analytics strategy with the business strategy. This misalignment leads to time, resources, and money being spent on projects that are not critical. And the costs of misalignment increase as these initiatives move from scoping phase to production phase. Furthermore, it is difficult to calculate the return on investment of a large data initiative without it meeting specific business goals. It is therefore a must that the organization's business strategy drives the data/analytics strategy.

For instance, at GSK in 2018, their data strategy was focused on near-term priorities and value creation. This in turn helped the firm establish the processes

and technical foundation necessary to bring about this transformation. This strategy, named "Value Strikes," was launched in 2018 and started yielding results in 2019 through a number of advanced analytics use cases. Jen Baxter, senior vice president, tech strategy and performance, says, *"The Value Strikes program was a way to accelerate our enterprise data and analytics ambition. Each use case, termed 'value strike,' leveraged our existing data in order to achieve our strategic priorities while delivering significant near-term P&L or cash value. In parallel, these experiences helped build up our people and technology capabilities across the organization."*[25]

Further, as organizations mature, they should ensure that each business unit's data and analytics strategy is in turn aligned with their own business unit's strategy. For example, if a business unit prioritizes growth over cost reduction, it should focus on data and analytics initiatives that bring in more sales whereas data governance and data quality initiatives, while are transversal, should also prioritize these objectives.

2 Build business users' trust in data

No matter how abundant your information, you will achieve little, if business users do not trust the data enough to use it for decision making. As seen in "Data trust is key to organizational agility, collaboration, and generating value," organizations where stakeholders trust enterprise data are more agile, collaborate better, and can monetize their data assets.

Establish data provenance and enable easy data discovery

Organizations today are able to collect and store large amounts of data from multiple sources, including numerous customer touch points, internal applications, social media and publicly available online information. In order to build trust, the following areas are key:

1. **Develop and maintain an enterprise data catalog**. Data catalogs are a collection of metadata, both technical and business, with data management and search capabilities. These catalogs should be able to give business users across different functions a standardized understanding of critical data assets. Organizations must also realize that meta data is not restricted to just technical meta data. It is even more important and probably more difficult to maintain the business metadata as this depends on the business context in which the particular data set has been generated and stored.

   Elaborating on their journey at Church & Dwight, Kevin Gokey, VP and global CIO – and Todd Retchless, director, data & analytics – say, *"We are not trying to boil the ocean in terms of cataloguing and defining data for the enterprise – we are doing it area by area. We have been focused in the past two*

years on the customer, our trade effectiveness and also on a supply chain front – this is where we think that there is a lot of opportunity to leverage data to drive synergies and take cost out of our operation."

2. **Maintain strong access controls**. Control on access might seem to be at loggerheads with the speed of access that the business requires. However, strong policies around data modification and deletion will ensure that data values remain uniform across the organization. While allowing "read" access to business users to enable data exploration, the technical teams must ensure that only the users from the right roles can modify data.

Develop guidelines on data ethics

Back in 2014, the team at Amazon started working on an AI-based program that can review the resumes of job applicants and select the top profiles. However, the company realized that the system's results were gender biased. Even after its efforts to keep the program gender-neutral, there was not any guarantee that the program would remain non-discriminatory and the company finally disbanded the program altogether.[26] Further, a study by the Centre for Data Ethics and Innovation – commissioned by the UK government – found that the country's police are concerned about the use of AI in their operations as it may lead to an increased bias. The fifty experts interviewed for this study which included senior police officers felt that AI may "amplify" prejudices.[27]

To build trust, and get people engaged with systems, organizations need to devise and enforce guidelines on the ethical use of AI. These guidelines should incorporate the key facets of explainability, interpretability, fairness, auditability, and transparency. For instance, HSBC has developed a set of global principles to address the Ethical Use of Big Data and AI. These principles aim to increase awareness of ethical implications, bring uniformity and predictability about using big data and AI, and to encourage colleagues to question and challenge before new use cases are adopted.[28]

We found that nearly 70% of data masters have processes to address the bias or fairness of algorithms and 60% use tools to help users in terms of "auditability" (see Figure 24).

Figure 24   **Business users at majority of data masters consider their AI algorithms fair and auditable**
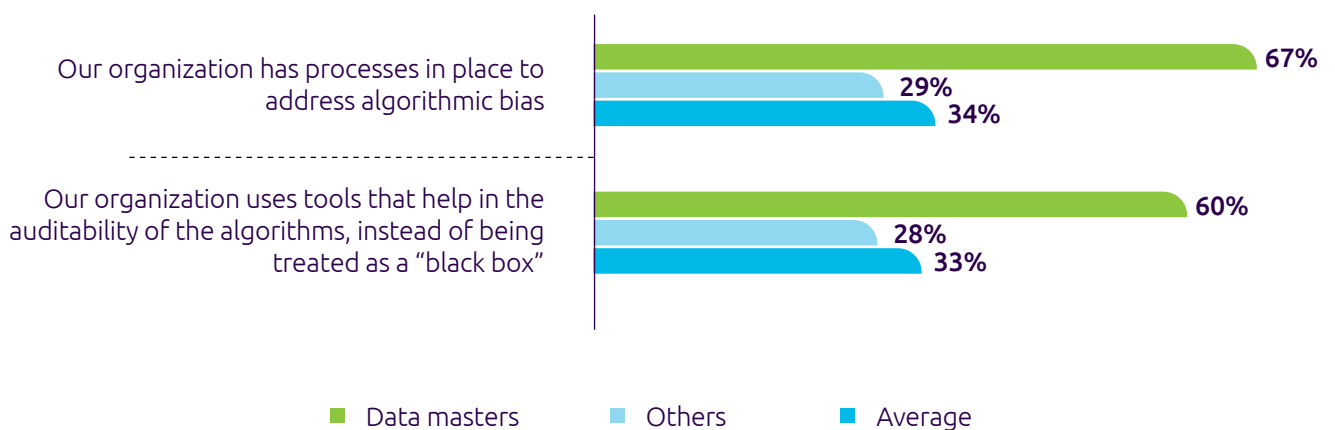
Our organization has processes in place to address algorithmic bias
- Data masters: **67%**
- Others: **29%**
- Average: **34%**

Our organization uses tools that help in the auditability of the algorithms, instead of being treated as a "black box"
- Data masters: **60%**
- Others: **28%**
- Average: **33%**

■ Data masters   ■ Others   ■ Average

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business unit executives, N=72 business data masters; N=432 business others.

3 Establish an AI and analytics CoE to assist business teams

To activate data and AI, AI and analytics CoEs need to be organized in a hub (enterprise) and spoke (line of business) structure. This federative way – balancing local ownership and a central platform drive – will be essential to infuse AI capabilities in each business unit. Ujjwal Sehgal, advance analytics leader at Mars Inc., says *"We started with a central team (hub) initially and delivered value, then slowly started fragmenting by market and by function (spoke). As we built these business-aligned teams, we positioned some of the central team members to fill those roles, which helped in quick adaption and scale to move at **100X speed**. As people move into this fragment, the core keeps getting replenished with newer team and they keep coming up to speed on what's happening across the enterprise as well."*

A CoE coordinates and steers data and AI activities across all business units, promotes innovation, and helps an organization to become a truly "data-powered." Bosch created the "Bosch Center for Artificial Intelligence" (BCAI) in 2017 to develop innovative AI technologies for the company. Using data from various business divisions, BCAI not only conducts research on AI, but also designs and implements AI for smart, connected, and autonomous technologies across Bosch's business sectors.[29]

This CoE can help an organization to:

• Identify the business use cases of AI and analytics
• Infuse AI tech throughout workflows and process
• Provide the resources and expertise needed for executing initiatives
• Develop skills required to become "data powered"
• Create a link between technology and business teams
• Break down the business KPIs into analytics capabilities
• Scale training in data-creation, model-building, and other AI-related processes across other facets of the business.

To align AI and analytics strategy with business strategy, it is also important to have a senior, strategic leader driving this goal. A chief AI/analytics officer provides development teams with a vision, establishes guidelines around prioritization of use cases, ethics, and security, and harmonizes the use of platforms and tools used for AI development. They should also work closely with the chief data officer. Our previous research on AI-powered enterprises highlights that more than 58% of AI-at-scale leaders have appointed an AI head/lead/chief AI officer (CAIO). [30]

4 Foster a data-powered culture by strengthening data citizenship

Understanding the inherent potential of data, and interpreting it properly, is crucial for strategic decision making by the business. We refer to this as "data citizenship" and these business users as "data citizens." They can explore and transform data and develop simple visualizations to highlight the insights using self-service analytical tools and without depending on the technical users. Further, they can also construct a narrative based on the data. It is critical today to stre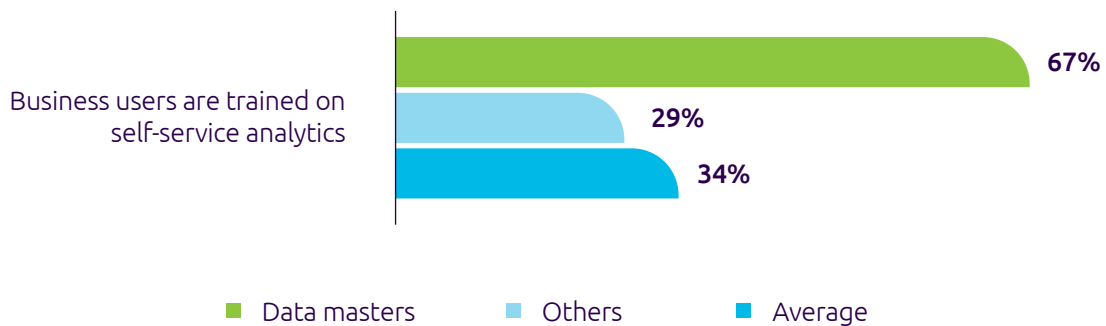ngthen this data citizenship in organizations. It is also critical to make data easily available to these data citizens while at the same time equipping them with required set of skills and tools necessary to perform this role. To drive data innovation, organizations are also looking at community engagement models such as code-a-thons, datathons, contests, etc. Our research shows that a majority of data masters have widespread data literacy (82%) and invest in a collaborative and innovation-driven data culture (75%). To build a data-first culture – and to arm employees with the analytical skills required to read, manipulate, analyze and argue with data – organizations need to train business users on self-service analytics tools and upskill users in areas such as data storytelling:

Train business users on self-service tools

With self-service analytics platforms, organizations can improve data citizenship across their business. Our research shows that 67% of leading organizations train their business users on self-service skills as compared to 24% others (see Figure 25). As Gokey and Retchless from Church & Dwight say,

"Self-service analytics is a big part of our strategy. Today, a lot of the data that has been analyzed throughout the organization is being manually prepped and moved. I think the paradigm shift that we see upcoming is when our data warehouse comes fully online and we have our curated data there for self-service analytics."

Figure 25 | **Data masters train their business users on self-service analytics**

Business users are trained on self-service analytics

- Data masters: 67%
- Others: 29%
- Average: 34%

■ Data masters   ■ Others   ■ Average

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business unit executives, N=72 business data masters; N=432 business others.

Lloyds of London, an insurance provider, focuses on ensuring that employees at all levels achieve more with data. With a user-friendly self-service website called "My Data," they provided their employees with an overview of all their business data, including access to apps and dashboards, and how data is used. The aim of this initiative is to improve data literacy and to enable business users to extract value from data.[31]

However, self-service implies that business users should be easily able to use the tools and the data. The tools used should serve the purpose for a data citizen – not a data scientist or a power user. And, training is a key enabler for self-service analytics. Joshua Merrill, principal reporting architect at Boston Children's Hospital, says, *"To be effective, a user needs to know what they're doing, and we've built a pretty big education department within [the hospital]."*[32]
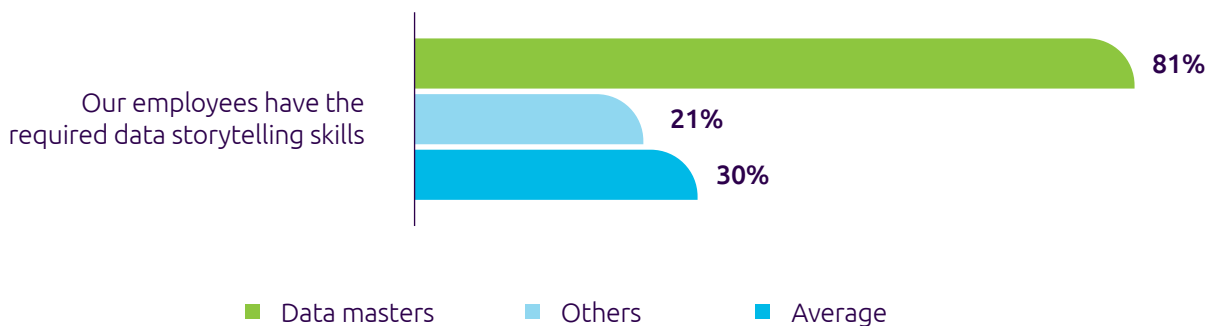
Improve on data storytelling skills

It is important to remember that information by itself means nothing unless it is understood and actioned by the right stakeholders. Organizations typically rely on dashboards to extract value from data. However, they are not always ideal as they often lack a compelling story or narrative. Too often, dashboards rely on a reader's ability to uncover insights and interpret information, which means they can introduce their own personal biases.

Data storytelling weaves data and visualizations into a tailor-made narrative for a specific audience. Narratives, visuals and data are merged together to explain, engage, and influence and drive change.[33] As Figure 26 shows, data masters are confident in their employees' storytelling skills. Singapore-based DBS bank has been working on equipping more than 16,000 employees across all its 18 Asian markets with big data and data analytics skills. Employees undergoing these programs are able to learn more about the concept of becoming a data-driven organization, IoT, data storytelling, the responsible use of data, and data visualization.[34]

Figure 26   Data masters are more confident their people possess storytelling skills

Our employees have the required data storytelling skills

81% — Data masters
21% — Others
30% — Average

■ Data masters   ■ Others   ■ Average

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=504 business unit executives, N=72 business data masters; N=432 business others.

81%

of data masters are confident in their employees' data storytelling skills.

Data enablers

1 Strengthen the data collection processes and improve data quality

Strengthen data collection:

Collecting operational data for existing processes, particularly for manufacturing industries, remains a challenge due to proliferation of devices and sensors, legacy systems, differences in communication technology and connectivity standards for the machines, etc. However, without data collection, plant managers will not be able to monitor the operations of the plants.

In order to improve the data collection processes, organizations must ensure:

- That all necessary machines are brought on to a network. For legacy equipment, this could mean developing a communication layer.

- That PLCs (programmable logic controls), I/O devices are integrated with production systems such as MES (manufacturing execution system), and ERP, etc. on the factory floor.
- That the data collection processes are designed by taking the inputs from the local teams or the operators working on a specific machine. This will avoid gaps that might otherwise remain hidden and will also help in making these operators understand the importance of feeding the right information into the systems.

Ensure data quality:

Data engineers and data scientists spend a lot of time on data cleansing and preparation. This necessary evil can be curbed to a certain extent by ensuring quality considerations are embedded into the setup of applications and processes. For instance, not providing a text box for user input when the field can only have a predefined set of values. All of the data masters we surveyed said that they consider data quality right from the design of applications and processes (see Figure 27).

Figure 27   **Data masters place a strong emphasis on data quality**
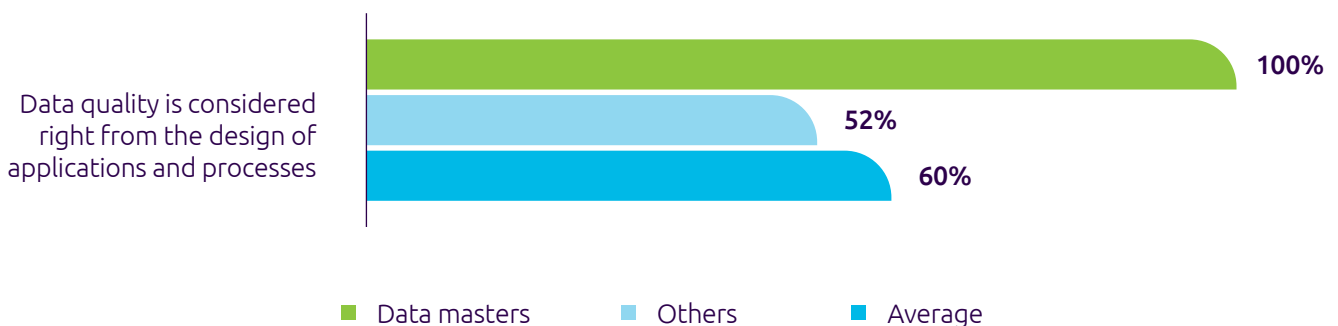
Data quality is considered right from the design of applications and processes

- 100% — Data masters
- 52% — Others
- 60% — Average

■ Data masters ■ Others ■ Average

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives, N=87 technology data masters; N=413 technology others.

Data quality is a continuous process and the rigor of addressing the quality issues needs to be adapted for each business unit. For instance, a marketing team might will not need the same level type of accuracy and data quality that an accounting team demands. Considering business constraints and specificities for each lines of business is critical in setting up the right data quality processes.
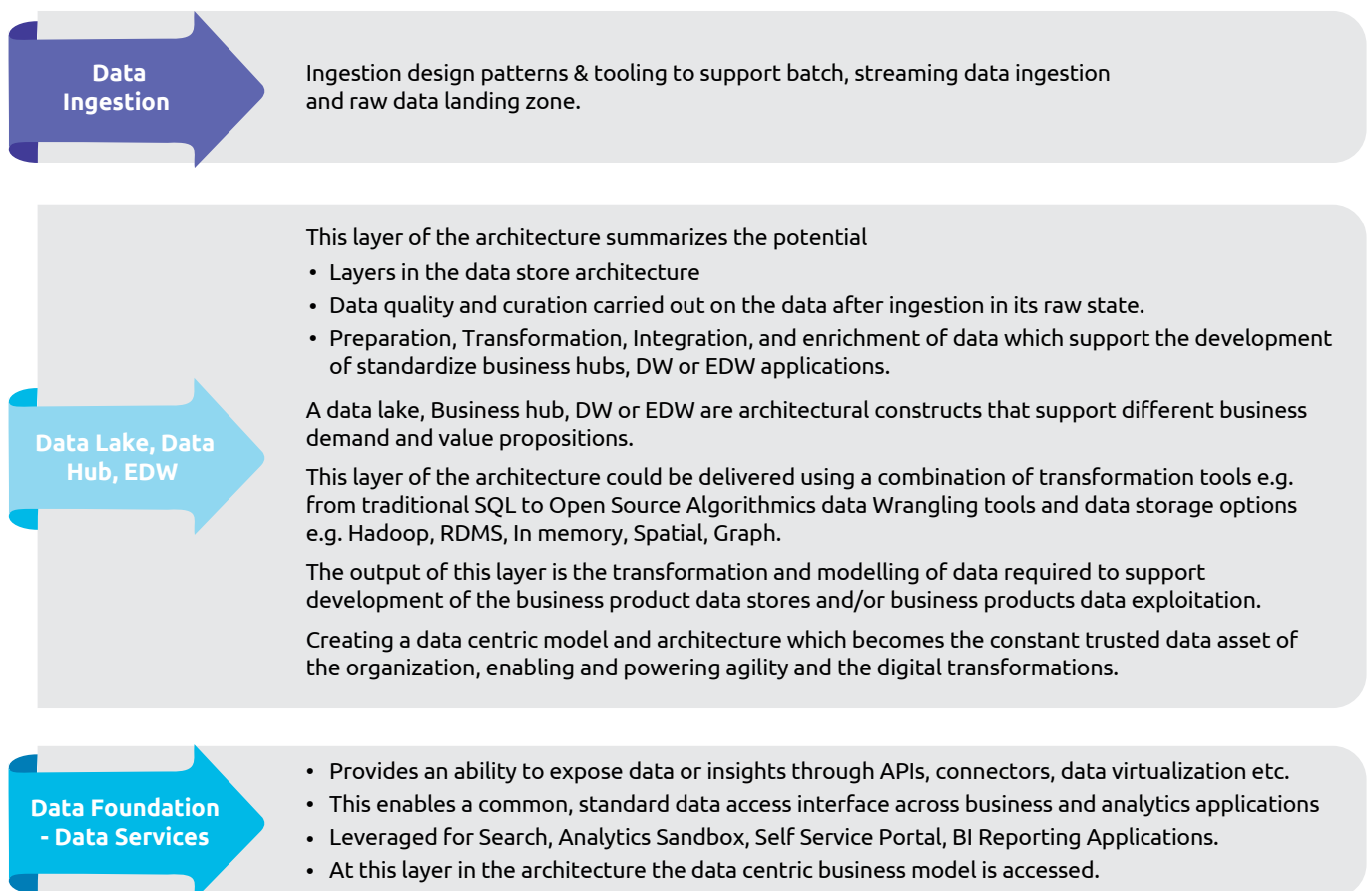
Invest in data landscape modernization to get agility in data activation

Monolithic architectures and legacy systems can compound the problem of data silos. Modernizing the data landscape is about supporting the organization's innovation and change by transforming the technology, and processes. The design of such data architecture should be guided by the following principles to ensure it is fit for the long-term future:

- Democratization of data – ability to provide the required data to the business users, at speed and in the right form and granularity, thereby allowing the business users to explore data and derive actionable insights without waiting for the technical teams.
- Data privacy and security
- Empowerment of business users – a platform that can support reporting, analytics and AI as well as self-service
- Scalability of service.

Components of a modern data estate include:

Figure 28 | **Data foundations of a modern data estate**

Data Ingestion

Ingestion design patterns & tooling to support batch, streaming data ingestion and raw data landing zone.

Data Lake, Data Hub, EDW

This layer of the architecture summarizes the potential
- Layers in the data store architecture
- Data quality and curation carried out on the data after ingestion in its raw state.
- Preparation, Transformation, Integration, and enrichment of data which support the development of standardize business hubs, DW or EDW applications.

A data lake, Business hub, DW or EDW are architectural constructs that support different business demand and value propositions.

This layer of the architecture could be delivered using a combination of transformation tools e.g. from traditional SQL to Open Source Algorithmics data Wrangling tools and data storage options e.g. Hadoop, RDMS, In memory, Spatial, Graph.

The output of this layer is the transformation and modelling of data required to support development of the business product data stores and/or business products data exploitation.

Creating a data centric model and architecture which becomes the constant trusted data asset of the organization, enabling and powering agility and the digital transformations.

Data Foundation - Data Services

- Provides an ability to expose data or insights through APIs, connectors, data virtualization etc.
- This enables a common, standard data access interface across business and analytics applications
- Leveraged for Search, Analytics Sandbox, Self Service Portal, BI Reporting Applications.
- At this layer in the architecture the data centric business model is accessed.

Source: Capgemini PerformAI/AI & Data Engineering.

Organizations are taking steps in this direction gradually. For instance, Uniper, an energy company based in Germany, created a data analytics platform with a central data lake in a cloud platform. This data lake gathers data from nearly 100 external and internal sources. The platform also has a data catalog with end to end metadata management capabilities.[35]

Our research shows that, at 84% of data masters, the analytics teams can cleanse data quickly (see Figure 29). A modern data estate as described above not only allows the analytics teams to access and prepare data quickly but also allows business teams to be more agile and responsive to changes in market dynamics.

Figure 29   **Analytics teams at the data masters spend less time in cleansing data**
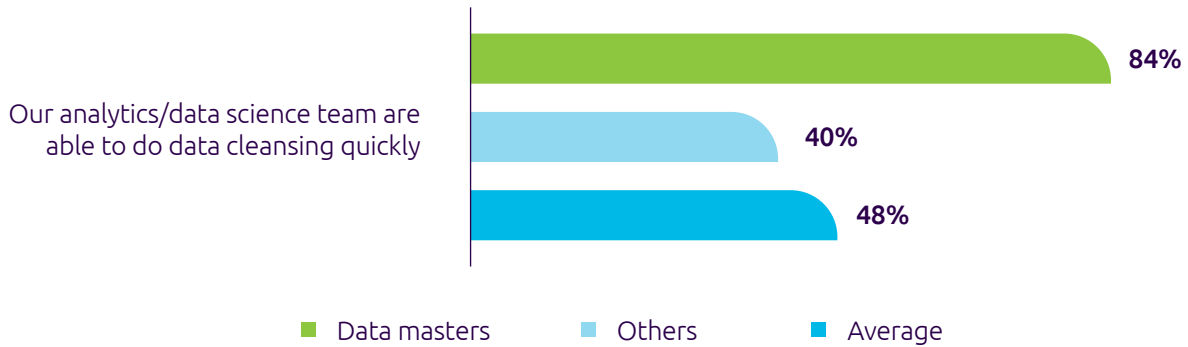
Our analytics/data science team are able to do data cleansing quickly

- Data masters — **84%**
- Others — **40%**
- Average — **48%**

■ Data masters   ■ Others   ■ Average

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives, N=87 technology data masters; N=413 technology others.
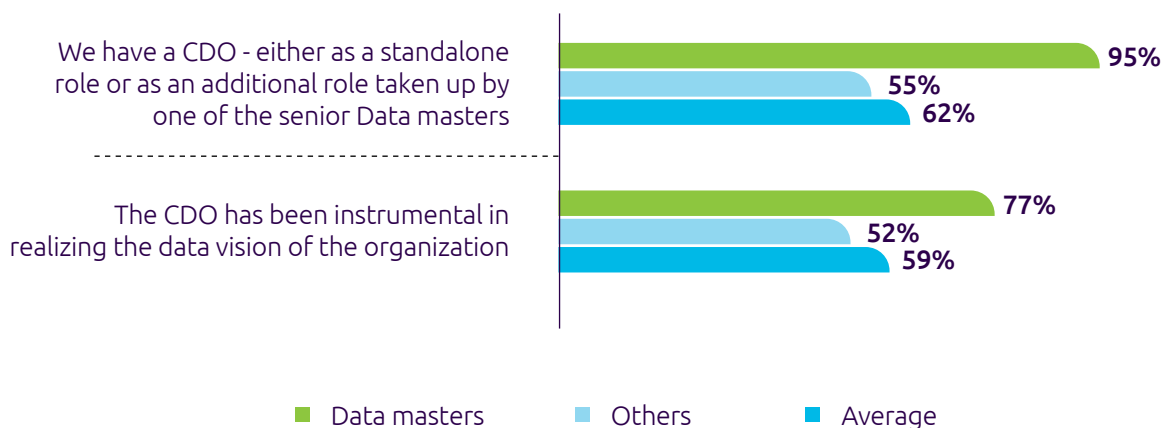
3 Operationalize data and analytics through DataOps and MLOps

Agile software development is well-suited to AI applications, where end users need to be involved early to help test, refine, and improve embedded algorithms. DataOps and MLOps extend the agile and DevOps practices from traditional software applications to data and AI-based applications. We define DataOps and MLOps as below:

- **DataOps** is a set of practices to democratize the use of data and improve its accessibility to business, by setting up an agile cooperative process with data analysts, data engineers, business users, and IT operations. It improves the quality, agility, speed of ingestion, and preparation as well as provisions data for use in AI and analytics use cases.
- **MLOps** is a set of practices to shorten the time to update and go live of analytics and self-learning systems. It improves the quality and robustness, by setting up a cooperative process involving data scientists, ML engineers, business users, and IT operations.

DataOps and MLOps are also part of the modern data estate (as described in Figure 28). DataOps and MLOps cover the entire cycle of gathering and preparing data for the models, running experiments, building and retraining models, and deploying and monitoring them. They can help an organization to:

- Catalyze the democratization of data.
- Monitor data "drift" and provide an early warning for retraining of models. This became very crucial during the changing world with COVID-19 pandemic. *"During times of economic distress and volatility, such as the pandemic, it is critical to halt some AI models or retrain them,"* says Apoorv Saxena, global head of AI technology at JPMorgan Chase.[36]
- Automate data operations resulting in fast, reliable, version-controlled, and better-quality applications for end-to-end orchestration of data
- Reduce the analytics cycle time and provide a faster access to actionable intelligence
- Provide an aggregated view over time of the entire dataflow, across the organization and out to end users
- Scale and industrialize data and AI to expand the benefits throughout the organization

In leading organizations, operationalizing insights is split and coordinated between DevOps, DataOps, and MLOps. As Figure 30 shows, 85% of data masters are deploying DataOps practices to improve the quality and speed of end-to-end data pipelines and 90% are using DataOps and MLOps to deliver analytical solutions quickly. Pharmaceutical firm, Celgene is using DataOps to automate data pipelines, monitor data quality, and automate deployment. With automation of their data pipeline, they were able to update 30X more visualizations per week and compress the cycle time required to produce new analytics from weeks to one day. This enabled the data analytics team to successfully address the volume of questions from sales and marketing, helping them maximize product adoption during the critical first phase of their product launch.[37]

We are using DataOps to deliver analytical solutions quickly
- **90%**
- **25%**
- **37%**

We employ DataOps to improve the quality and speed of end-to-end data pipelines including automation
- **85%**
- **27%**
- **37%**

■ Data masters   ■ Others   ■ Average
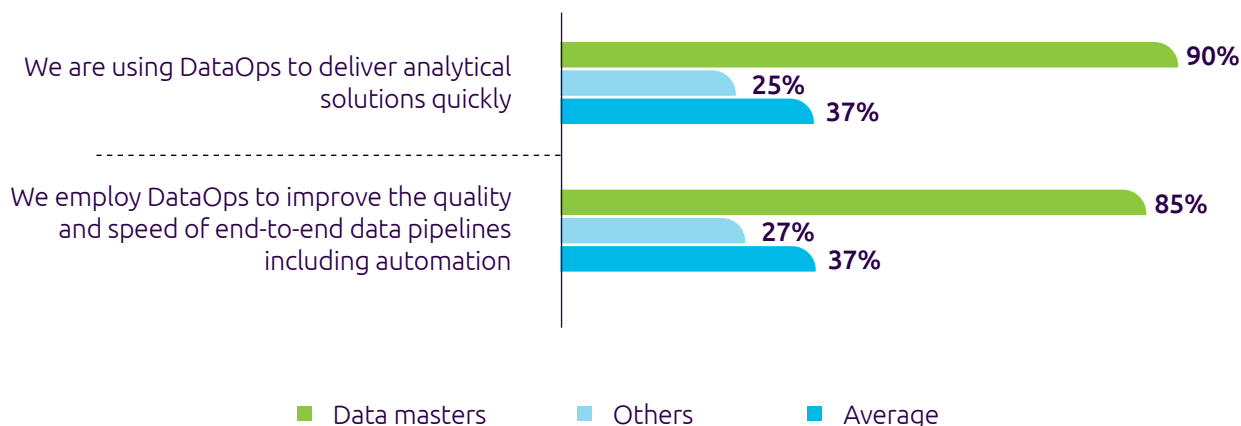
Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives, N=87 technology data masters; N=413 technology others.

4   Adapt the data governance as data mastery evolves

A chief data officer helps in setting the vision and drives the governance

The data masters we surveyed have one thing in common – 95% of them have a chief data officer (CDO), either as a standalone role (87%) or as an additional responsibility (8%). And as Figure 31 shows, a majority of them believe that the CDO has been instrumental in realizing their data vision.

We have a CDO - either as a standalone role or as an additional role taken up by one of the senior Data masters
- **95%**
- **55%**
- **62%**

The CDO has been instrumental in realizing the data vision of the organization
- **77%**
- **52%**
- **59%**

■ Data masters   ■ Others   ■ Average

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives, N=87 technology data masters; N=413 technology others.

A CDO's responsibilities include:

- **Strategy and roadmap**: Managing and facilitating the organization's ability to meet the business objectives
- **Governance**: Establishing the operating model and driving data governance
- **Operationalization of data/analytical initiatives**: Playing a hands-on role or an overseeing role based on the data mastery of the organization in the implementation of the initiatives
- **Guiding principles**: Developing principles around data privacy
- **Data quality**: Enriching the data quality through the right design and policy framework
- **Data-driven practices**: Driving a culture of data-driven innovation and establishing a data-driven culture
- **Talent**: Working together with HR/Learning & Development to nurture the skillsets required to make the organization data powered
- **Monetization**: Supporting the organization's transformation by exploring the potential of monetizing internal and external data.

Depending on the organization, a CDO's role also includes:

- **Data estate modernization**: Architecting the modernization of the data landscape as seen in the previous section
- **Data and AI ethics**: Establishing ethical guidelines and standards around data and analytics. For instance, data ethics is a chief component of her CDO role at Mastercard for JoAnn Stonier.[38]

Clearly, CDO must wear multiple hats in an organization. Given the prominence of the role, it is perhaps not surprising that in 36% of organizations we surveyed, the CDO directly reports to the CEO. In another 48% of the organizations, the CDO reports to the information management leadership roles such as the CIO/CTO, or the Chief AI Officer. A CEO line of control helps organizations in backing the data/analytical initiatives and works in making the CDO's role more effective. To effectively handle the role, a CDO must also work closely with CIOs/CTOs to collaborate on the technical operations, policies for access and usage as well as with Business CXOs (COO, CMO, CFO, etc.) to understand the business strategy and steer the data strategy in accordance with the business strategy.

Customize the data governance as per the mastery of the organization

At many data laggards, data initiatives are handled in silos by different business teams themselves. There is also an absence of an enterprise data governance and policies on data privacy, ethics, and even sustainability. Such a siloed approach can result in little data standardization, and no enterprise-level view for the C-level executives.

We believe that the operating model for data governance should evolve from a centralized model to a hub-and-spoke model where a CDO-led central unit will be responsible for policy making and governance while the operationalization of initiatives is managed by the business units themselves. A CDO's role here is to provide guidance and direction to the organization while allowing the spokes to operate independently.

Some organizations also establish a data center of excellence (COE) to help their analytics and AI teams. For instance, Morgan Stanley has set up a data CEO with a team of 30 experts from data architecture, infrastructure, and governance, partly to ensure AI and the other applications are built with the right data. *"We want to make sure we have high-quality data, because your insights are only as good as the data that underpins it,"* said Gez Hester, head of the Data Center of Excellence. *"High-quality data is data that's accurate, up-to-date and complete. To achieve that, the data team must catalogue all the data that's available in, for example, the commercial real-estate loan business and understand the data's provenance, where it is being sourced from, and how it is going to be used."*[39]

Ultimately, the right operating model depends on the organizational context (i.e., structure, culture, processes). So, in organizations with a decentralized culture, a bottom-up approach where various teams independently scope, develop and execute their solutions could also work best.[40]
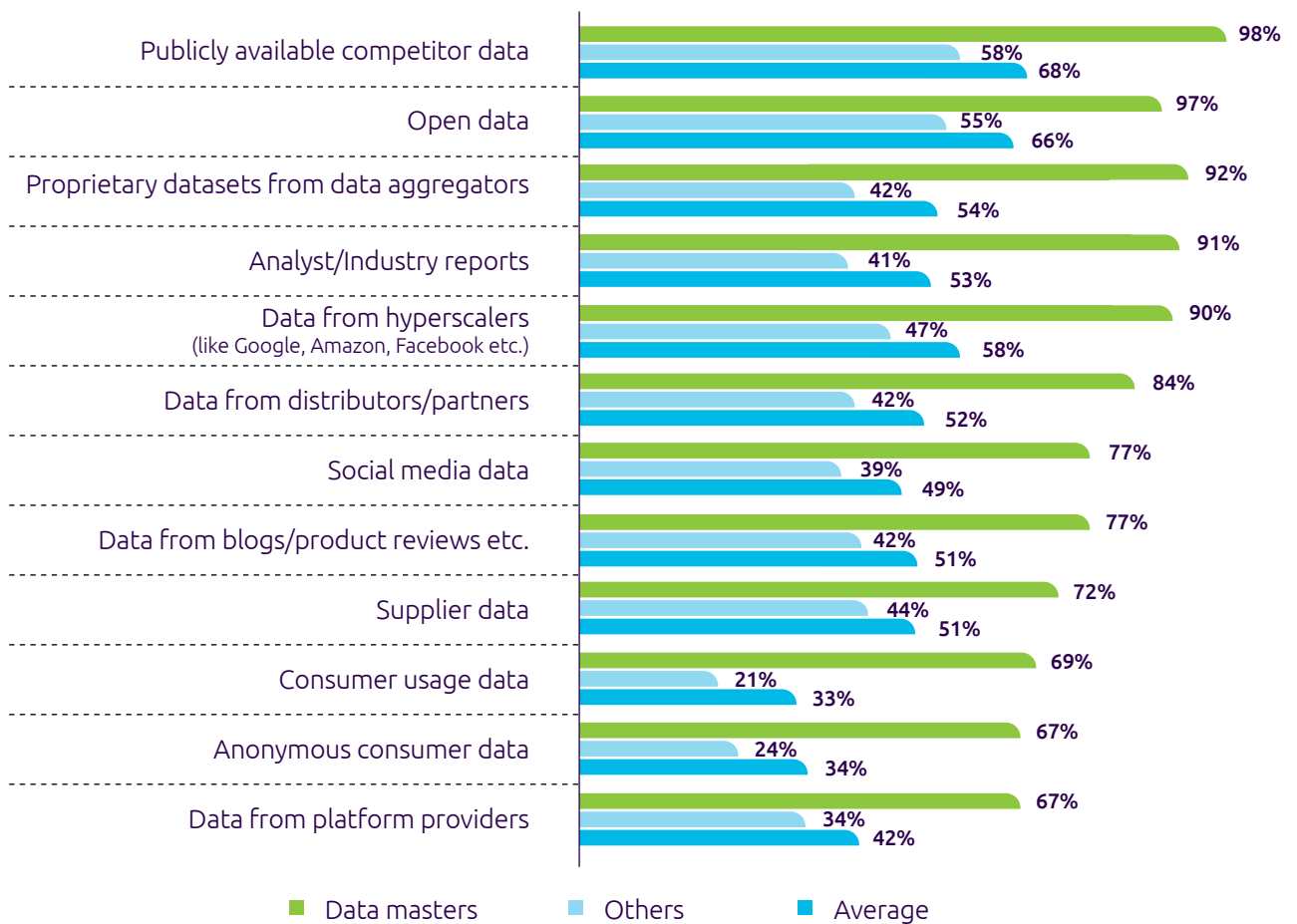
Leverage external data to enhance insights

Data masters look at unique sources of information – including external – that give them a competitive edge. External data analysis can help organizations to evaluate the risks and opportunities that they would miss with inputs limited to data generated from internal operations, customers, and first-tier suppliers. Supermarket company Tesco, for example, uses weather data to drive richer insights that help them to predict sales and stock requirements.[41]

Data-driven partnerships with clients, suppliers, and other partners – such as data disruptors – are becoming increasingly critical for monetization of data and for enhancing the business value of insights.

Our research reveals that a majority of leaders do not limit their data sourcing to traditional external sources – such as open data. They also go beyond and leverage non-traditional sources such as data-aggregators (such as Nielson, Experian), or data disruptors (such as Google, Amazon, Facebook) or data from distributors/partners and clients (see Figure 32).

Figure 32   **Data masters leverage variety of external data sources to enhance the insights**

Source	Data masters	Others	Average
Publicly available competitor data	98%	58%	68%
Open data	97%	55%	66%
Proprietary datasets from data aggregators	92%	42%	54%
Analyst/Industry reports	91%	41%	53%
Data from hyperscalers (like Google, Amazon, Facebook etc.)	90%	47%	58%
Data from distributors/partners	84%	42%	52%
Social media data	77%	39%	49%
Data from blogs/product reviews etc.	77%	42%	51%
Supplier data	72%	44%	51%
Consumer usage data	69%	21%	33%
Anonymous consumer data	67%	24%	34%
Data from platform providers	67%	34%	42%

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=500 technology executives, N=87 technology data masters; N=413 technology others.

The size and complexity of the external data sources are a challenge for organizations to fully leverage the potential. Moreover, data preprocessing and cleansing is time consuming yet critical. Some best practices emerge for utilizing external data:

- Define which external data sets will allow you to change the game and to enhance your customer, process, market, or ecosystem insights.
- Create catalogs of external data: Although it requires updating information management processes that were typically designed to handle only internal data, it is worth the effort. A data catalog will reveal what data the organization has access to, where it resides, who is responsible for it, and other key metadata required to generate value from the data.
- Create specialist roles for external data sourcing: For instance, at European reinsurance provider, Munich Re, data team includes 20 "data hunters" – who are dedicated to data sourcing.[42]
- Introduce a governance framework for oversight of external data sourcing and acquisition which comprises of cross-functional data owners.
- Build "data collaboration" spaces with partners, clients, suppliers in order to collaborate and get insights on data sets sitting outside your own organization.
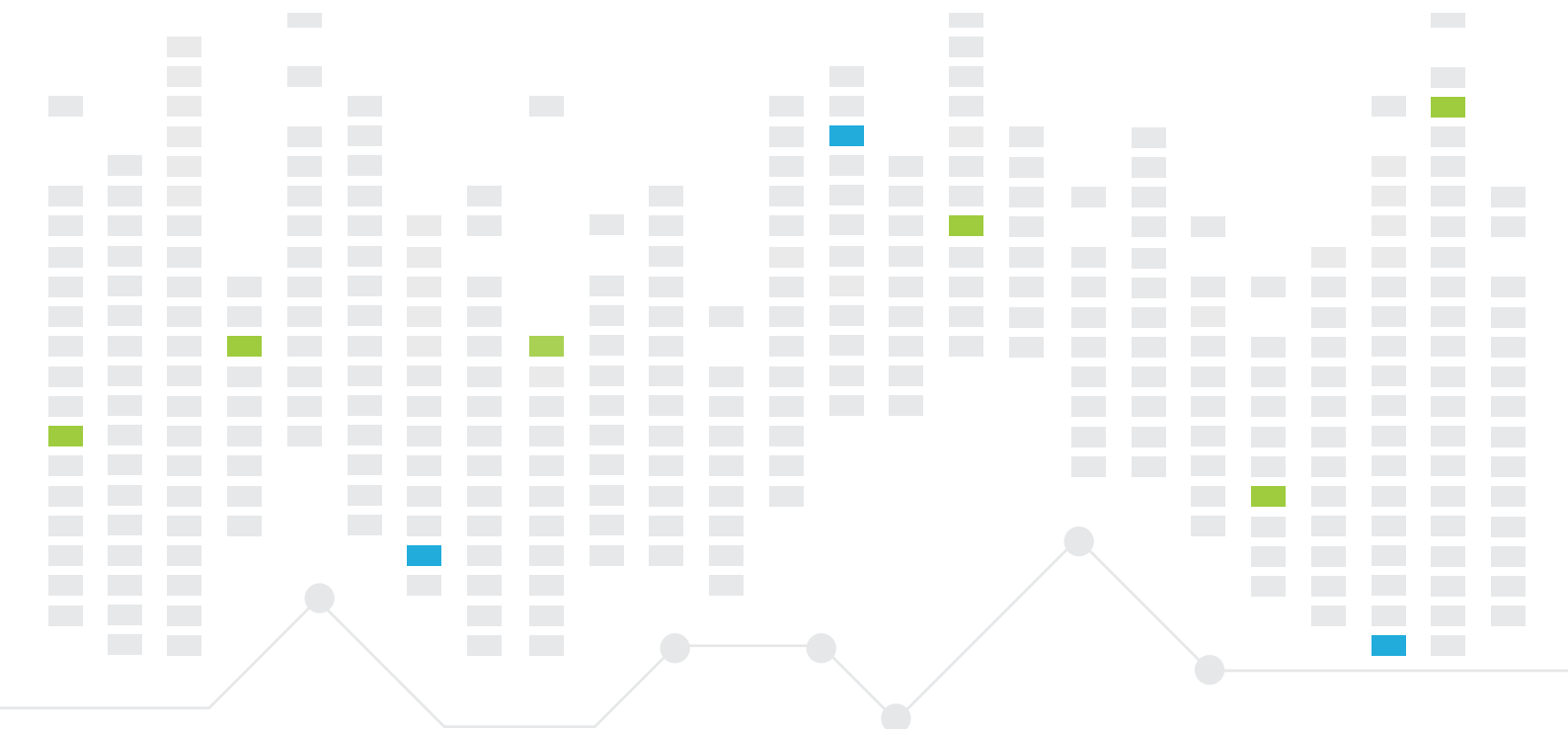
Conclusion

As we have seen, the world is producing more and more data. However, we have also seen that proliferation of data does not necessarily turn into increased organizational intelligence. More and more organizations say they are data powered, but we uncovered a significant number of organizations that are not able to build forward-looking insights, monetize their data, or build any sort of competitive advantage on the back of higher data volumes.

We have covered a range of challenges that are standing between organizations' hopes for data and the reality. But one of the most striking is lack of trust. This is a major problem. If you cannot trust the data, how is an executive ever going to trust in the outcome of that data? As we all know, trust takes a lifetime to build. Leaders need to start focusing now on building trust into their data strategy and approach, as without it, moving from data to value creation will prove to be impossible.

The financial performance of data masters clearly stresses on the need for organizations to improve upon the aspects of their data mastery. Further delay in taking necessary actions will only increase the data divide. In addition to investing in the foundations and the behaviors necessary, organizations must work towards activating data, i.e., transforming the data and insights into business outcomes. Finally, the truly data-powered organizations make extensive use of data available to them from various external sources and collaborate with customers, suppliers, particularly hyperscalers to gain an edge.

Appendix 1

Illustrative list of use cases examples for each business objective

Business objective	List of use cases
Customer engagement	• **Brand sentiment**: Understanding brand sentiment helps in measuring the success of a marketing campaign, improving product quality, identifying influencers, etc. By reviewing the social data from social media channels such as Twitter, Facebook, and Instagram, of the brand as well as its competing brands, organizations can monitor their brand health. • **Psychographics personalization**: Psychographics is a qualitative methodology used to describe consumers on psychological attributes. By capturing the user behavior through their social media profile, an AI-based psychographic personalization experience can be implemented.
Operational efficiency	• **Predictive maintenance**: This use case captures and analyzes real-time information about root causes of equipment problems, enabling manufacturers to reduce unplanned downtime and defects, anticipate machine failure, and predict future maintenance needs. By capturing, integrating and analyzing machine data with predictive maintenance, operators gain visibility into machine performance. Proofs of concept have shown that the lifetime extension of aging assets was increased up to 15–30%, uptime improvement between 20–40%. A project carried for the biggest APAC plant of an Automotive OEM led to 500mn/week of downtime saved for 600 robots • **Automatic product quality detection**: QA (quality assurance) processes still are not fully automated and are done manually. This results in non-uniform inspection which is also prone to fatigue. AI solutions powered by the right data can improve the QA process with a 99.7% accuracy and can work on applications such as checking tube installation, proper part coloring, paint job quality, etc.
Cost savings	• **Minimize production loss due to machine stoppages**: Stoppages can happen due to multiple reasons like technical failures within the machine, material not available in time, external factors (electricity, absenteeism, etc.). Using machine learning to analyze the historical reasons, manufacturers can predict the machine and the line that would most likely fail and take proactive actions on these stations to minimize loss of production. • **Identifying fraudulent transactions at point of purchase**: Manual reviews of transactions ex post facto is costly, time-consuming. Machine learning can be used to predict the likelihood of a fraud to happen in the future. The system can deliver the confidence levels required to approve or decline a transaction, in milliseconds and are more effective than humans at detecting subtle or non-intuitive patterns to help identify fraudulent transactions.
Top-line benefits	• **Dynamic price management**: With the price comparisons available to customers at the touch of a button, retailers should be able to change prices of the items in store as well as online based on factors that could include time of day, demand, location, competition, and customer buying patterns. By using historical data and external data and determining the elasticity of pricing, the optimal price can be determined. This will help retailers in improved demand forecasting accuracy and increased revenue.

Appendix 2

Statements considered for the technology executives model

Category	Theme	Statement
Data foundations (tech & tools)	**Data & AI platform**	We have invested in analytics tools and platforms
		We have dedicated resources for retraining and redeployment of existing models
	Data identification	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
		Our organization has a complete picture of all the data inventory
		We have a record of the data owner(s) for all internal data
	Data ingestion, processing, and harvesting	We are able to leverage structured, semi-structured as well as unstructured data for decision making
		We have defined and continue to update an enterprise-wide data catalog to assist data users
		We have automated the process of data collection
		We can store, retrieve and analyze the data at all levels of the organization's value chain
		We have developed data management processes that source, clean, prepare, integrate, and provide access to data at the speed that the business needs
		We have deployed processes to check quality of data at the points of capture
		We ensure data stewardship in each data domain for improving data quality
	Data governance implementation	Our data governance considers and plans for varying maturity levels of each business unit
	Infusion into business for data activation	We are expanding data, BI, and analytics in the cloud
		Higher-than-average number of our data sources are in cloud
		We offer data preparation tools for self-service data management
	"Data advantage" tools to leverage external data	We use the following external data for decision making – supplier data, data from distributors/retailers, data from platform providers, anonymous consumer data (such as cookies), personal identifiable information (PII) of consumers, consumer usage data, social media listening data, data from blogs/product reviews, publicly available competitor data, analyst/industry reports/reports from agencies such as D&B, data from hyperscalers such as Google, Amazon, Facebook, proprietary datasets from data aggregators such as Nielsen, Experian, open data, and other external data (such as weather, traffic)
		We create data by designing products and processes to capture new data

Category	Theme	Statement
Data behaviors	Data-powered decision making in business	We use predictive, prescriptive, autonomous/self-optimizing approaches for decision making in our organization
		Business teams work with IT/data teams to identify insights from the data troves
	Data governance processes	We build cross-functional data and insights teams that work with data engineers, data scientists, solution architects, and software developers
	Data activation culture	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
	Data guiding principles	Data is FAIR (Findable, Accessible, Interoperable, Reusable), secured, and sustainable
		Executives leaders champion data privacy, security, and ethics and are accountable for success
	Data advantage strategy	We monetize data assets/insights through our products and services
		We continually act on insights to optimize business outcomes
	Data activation vision and strategy	We have role-based data upskilling programs for most of our employees
		We upskill our employees on data skills such as model training, course correction, and maintenance in addition to their regular AI/ML skillset
		We train our business users on analytical and storytelling skills
		Our senior leadership is fully committed towards appropriate investments in resources and technology to make us data powered
		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

Statements considered for the business executives model

Category	Theme	Statement
Data foundations (tech & tools)	Data & AI platform	We have invested in analytics tools and platforms
		Our organization has dedicated resources for retraining and redeployment of existing models
	Data identification	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

Category	Theme	Statement
Data foundations (tech & tools)	Data ingestion, processing, and harvesting	We are able to leverage structured, semi-structured as well as unstructured data for decision making
		Our data catalogue is clearly defined and regularly updated for us to use
		Our data collection process is automated
		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
	Data governance implementation	Our data governance considers and plans for varying maturity levels of each business unit
	Infusion into business for data activation	We have expanded data, BI, and analytics in the cloud
	"Data advantage" tools to leverage external data	We use the following external data for decision making – supplier data, data from distributors/retailers, data from platform providers, anonymous consumer data (such as cookies), personal identifiable information (PII) of consumers, consumer usage data, social media listening data, data from blogs/product reviews, publicly available competitor data, analyst/industry reports/ reports from agencies such as D&B, data from hyperscalers such as Google, Amazon, Facebook, proprietary datasets from data aggregators such as Nielsen, Experian, open data and other external data (such as weather, traffic)
		Our products and processes are designed to capture new data

Category	Theme	Statement
Data behaviors	Data-powered decision making in business	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
		Business teams work with IT/data teams to identify insights from the data troves
	Data governance processes	We build cross-functional data and insights teams that work with data engineers, data scientists, solution architects, and software developers
	Data activation culture	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

Category	Theme	Statement
Data behaviors	Data guiding principles	Sustainability (data production, storage, and use are sustainable) of data is a key consideration in our data governance
		We actively promote the exploration, collaboration of new ideas, and experimentation at all levels
		Executives leaders champion data privacy, security, and ethics and are accountable for success
	Data advantage strategy	We monetize data assets/insights through our products and services
		We continually act on insights to optimize business outcomes
	Data activation vision and strategy	Our employees realize the value of being data driven
		We have role-based data upskilling programs for most of our employees
		We are trained on analytical and storytelling skills
		Our senior leadership is fully committed towards appropriate investments in resources and technology to make us data powered
		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

Research Methodology

In this research:

- We surveyed technical executives working in IT/data functions such as IT, information management, data management and analytics, business intelligence, etc. from 500 organizations.
- We surveyed business executives working in business functions from 504 organizations.

Both surveys were conducted in August of 2020 and covered organizations with at least USD 1 billion in revenues for the previous financial year. We further conducted a financial analysis of 700+ unique organizations from the above set, for the financial years 2018–19 and 2019–20.

In addition, we conducted more than 15 in-depth interviews with senior technical and business interviews on the data initiatives in their organizations.

Distribution of organizations in the market surveys

Organizations by country

7% 7% 7% 10% 10% 8% 8% 7% 8% 8% 10% 10%

- 🇦🇺 Australia
- 🇨🇳 China
- 🇫🇷 France
- 🇩🇪 Germany
- 🇮🇹 Italy
- 🇮🇳 India
- 🇳🇱 Netherlands
- 🇪🇸 Spain
- 🇸🇪 Sweden
- 🇬🇧 UK
- 🇺🇸 US
- 🇸🇬 Singapore
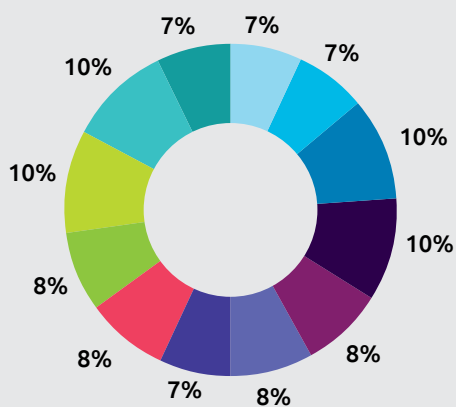
Organizations by industry

10% 10% 10% 10% 10% 10% 10% 10% 10% 10%

- Consumer Products Manufacturing
- Retail
- Banking
- Industrial Manufacturing
- Automotive
- Telecom
- Public Services
- Energy & Utilities
- Life Sciences & Healthcare
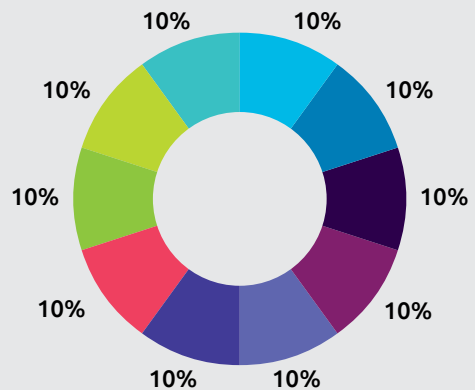
Organizations by revenue

- 51%
- 30%
- 13%
- 6%
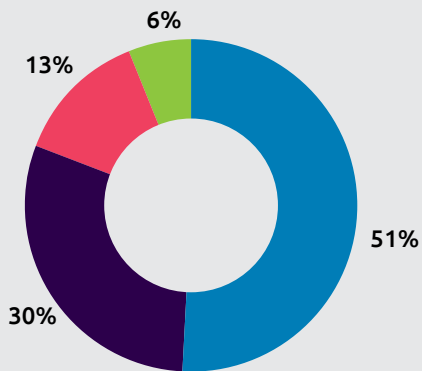
- ■ USD 1 billion - less than USD 5 billion
- ■ USD 5 billion - less than USD 10 billion
- ■ USD 10 billion - less than USD 20 billion
- ■ More than USD 20 billion

Respondents by designation

- 1%
- 11%
- 36%
- 52%
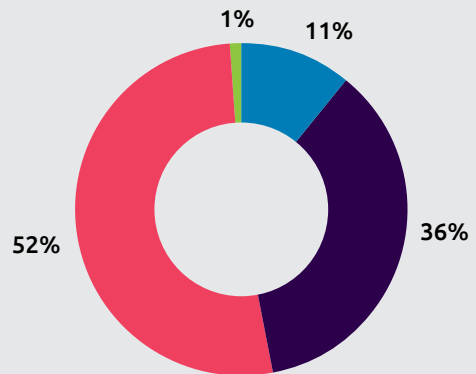
- ■ Executive (C-Level)
- ■ Vice President
- ■ Directors
- ■ Partner/Board Member

Technology executives by function

- 19%
- 9%
- 7%
- 9%
- 7%
- 9%
- 7%
- 16%
- 7%
- 10%
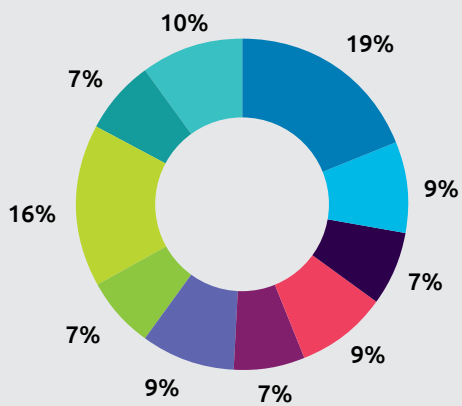
- ■ Information technology /Information management
- ■ Digital
- ■ Information security
- ■ Data management and insights
- ■ Data architecture
- ■ IT/Data Innovation
- ■ AI and Analytics
- ■ Business intelligence
- ■ Data warehousing
- ■ Data governance

Business executives by function

- 8%
- 10%
- 15%
- 7%
- 8%
- 16%
- 6%
- 8%
- 13%
- 9%
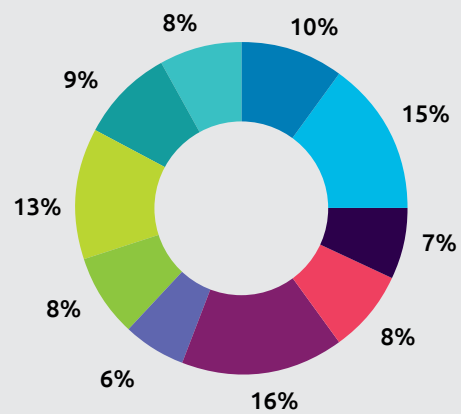
- ■ General management /Strategy
- ■ Sales andmarketing
- ■ R&D/Innovation
- ■ Production/manufacturing
- ■ Operations
- ■ Procurement/Supply chain
- ■ Customer service
- ■ Finance and accounting
- ■ Risk and Compliance
- ■ Human Resources

Source: Capgemini Research Institute, Data-powered enterprises survey, August 2020, N=1,004 organizations, N=500 technology executives, N=504 business unit executives.

References

1. Guo, Huadong & Wang, Lizhe & Chen, Fang & Liang, Dong. (2014). Scientific big data and Digital Earth. Chinese Science Bulletin (Chinese Version). 59. 1047. 10.1360/972013-1054.
2. Capgemini Research Institute, "Understanding digital mastery today: Why companies are struggling with their digital transformations," July 2018.
3. Computer Weekly, "How Dell is leveraging data to reopen its offices," August 2020.
4. NS Healthcare, "What do NHS trusts in the UK need to consider when rolling out data strategies?" August 2020.
5. Computer Weekly, "How Bank Mandiri is driving the use of big data analytics," September 2020.
6. Businessofgovernment.org, "Implementing the DATA act of 2014: Insights from Christina Ho, deputy assistant secretary, accounting policy and financial transparency, U.S. Department of the Treasury," accessed October 28, 2020.
7. Forrester, "Determine your data's worth: Data plus use equals value," February 2019.
8. Reuters, "Dollars in the detail: banks pan for gold in 'data lakes'," June 2019.
9. Detroit Free Press, "Data could be what Ford sells next as it looks for new revenue," November 2018. Jim Hackett stepped down as the president and the CEO of Ford and was succeeded by Jim Farley on October 1, 2020.
10. Tech Target, "3 companies successfully monetizing IoT data," May 2019.
11. IoT world today, "A look at Michelin's product-as-a-service strategy," February 2020.
12. Forrester, "The rise of the subscription business model," April 2020.
13. Wired, "How data helps deliver your dinner on time—and warm," February 2019.
14. MIT Sloan Management Review, "Seizing opportunity in data quality," November 2017.
15. Capgemini, "How data quality can hurt your data science programme… if you're not careful," February 2020.
16. Capgemini Research Institute, "The AI-powered enterprise: Unlocking the potential of AI at scale," July 2020.
17. Capgemini Research Institute, "Smart factories@scale: Seizing the trillion-dollar prize through efficiency by design and closed-loop operations," November 2019.
18. Forbes, "10 examples of predictive customer experience outcomes powered by AI," December 2018.
19. Barrons, "Nike buys another analytics firm as it dives into metrics on consumer behavior," August 2019.
20. Capgemini Research Institute, "The AI-powered enterprise: Unlocking the potential of AI at scale," June 2020.
21. The Wall Street Journal Pro, "CIO Helps Conagra Turn Food Trends Into Products," July 22, 2019.
22. Forbes, "Big data at Volvo: predictive, machine-learning-enabled analytics across petabyte-scale datasets," July 2016.
23. Teradata, "Danske bank fights fraud with deep learning and AI," 2018.
24. Capgemini, "Capgemini announces Project FARM," October 2019.
25. CIO, "GSK takes targeted approach to data-driven transformation," July 2020.
26. Reuters, "Amazon scraps secret AI recruiting tool that showed bias against women," October 2018.
27. AI News, "UK police are concerned AI will lead to bias and over-reliance on automation," September 2019.
28. HSBC. (n.d.). *Our conduct*. Retrieved October 1, 2020 from https://www.hsbc.com/our-approach/risk-and-responsibility/our-conduct
29. Bosch, "Bosch Center for Artificial Intelligence". Retrieved October 5, 2020 from https://www.bosch-ai.com/.
30. Capgemini Research Institute, "The AI-powered enterprise: Unlocking the potential of AI at scale," July 2020.
31. Vizlib, "Data literacy: Closing the gap and opening up new business opportunities," February 2020.
32. Tech Target, "Success of self-service BI depends on solid infrastructure," August 2020.
33. Forbes, "Data Storytelling: The Essential Data Science Skill Everyone Needs," November 2016.
34. Strategy Focused Group, "DBS Bank equips 16,000 employees with data skills for the future," November 2019.
35. Talend, "Talend Enables Energy Company Uniper to Analyze Data at Scale in the Cloud," December 2018.
36. The WSJ Pro, "Pandemic-related disruptions hurt reliability of some AI models; Humana and JPMorgan are among the companies watching for AI 'model drift' as lockdowns and social-distancing rules bring about unforeseen trends," June 10, 2020.
37. DataKitchen. (n.d.) Celgene DataOps case study: meeting the product launch challenge with DataOps.
38. Harvard Business Review, "Are you asking too much of your chief data officer?" February, 2020.
39. The Wall Street Journal, "Morgan Stanley Center of Excellence Readies Bank for AI's Data Demands; Companies looking to supercharge decision making through AI first need to prepare troves of data," April 2019.
40. Capgemini Research Institute, "The AI-powered enterprise: Unlocking the potential of AI at scale," June 2020.
41. Innovation Enterprise, "Are companies using their external data?" accessed October 29, 2020.
42. Forrester, "The insights professional's guide to external data sourcing," May 2019.

About the Authors

Anne-Laure Thieullent
Managing Director, Artificial Intelligence &
Analytics Group Offer Leader
annelaure.thieullent@capgemini.com

Anne-Laure leads the Artificial Intelligence &
Analytics Capgemini Group Offer (Perform AI),
one of Capgemini's 7 Group Portfolio Priorities.
She advises Capgemini clients on how they should
put AI technologies to work for their organization,
with trusted AI at scale services for business
transformation and innovation. She has over 19
years of experience in massive data, analytics and
AI systems.

Valérie PERHIRIN
Managing Director, Insight driven enterprise
Capgemini Invent
valerie.perhirin@capgemini.com

Valérie leads the data, analytics and artificial
Intelligence consulting activities. She advises
Capgemini Invent clients to design their data
strategy and designs roadmap and execution plan.
She helps our clients to drive their data powered
transformation in order to reach operational
excellence and leverage data and AI to invent
new sources of revenue. She has over 19 years of
experience in data centric transformation powered
by data, analytics and AI.

Steve Jones
EVP, Chief Data Architect
steve.g.jones@capgemini.com

Steve is the original author of the industries first
standard on Big, Fast, Managed Data (the Business
Data Lake) and has been working on cloud and
big data technologies for over a decade. He has
been helping organizations adopt new approaches
while simplifying and rationalizing their technology
infrastructure.

Conor McGovern
Vice President, Insight Driven Enterprise,
Capgemini Invent
conor.mcgovern@capgemini.com

Conor McGovern leads the Insight Driven
Enterprise (IDE) practice in Capgemini Invent UK.
Conor and his team use data, analytics and AI to
tackle the toughest business challenges for clients.
They help drive strategic, real-time decision-
making, eliminate repetitive tasks and enable
new levels of efficiency. His 20+ year career has
been spent in consulting, analytics and business
development.

Zhiwei Jiang
CEO, Insights and Data Global Business Line
zhiwei.jiang@capgemini.com

Zhiwei is the CEO of I&D Global Business Line for
the last two years. Before that, Zhiwei was the
Head of Insights & Data for Financial Services
Business Unit. Prior to Capgemini, he spent 20
years in banking. His passion has been in risk,
regulatory, data privacy, security and now more
into AI ethics and social impact.

Ron Tolido
EVP, CTO, Chief Innovation Officer, Insights and
Data global business line
ron.tolido@capgemini.com

EVP, CTO and Chief Innovation Officer, Capgemini
Insights & Data global business line. AI Futures
domain lead within Capgemini's Technology,
Innovation & Ventures (TIV) council. Lead author of
Capgemini's TechnoVision trend series. Certified
Master Architect.

Thomas Svahn
Vice President, Deputy MD I&D Scandinavia, Head
of Insights & Data Sweden | Advectas, part of
Capgemini | Insights & Data
thomas.svahn@capgemini.com
thomas.svahn@advectas.se

Thomas is a strategic advisor focusing on analytics
strategy and connecting Analytics to KPIs and
business outcomes. He has over 20 years of
experience from Business Intelligence, Data
Management and advisory services. He has a
Master of Business Administration from University
of Gothenburg, Sweden.

Nick Clarke
Head of Analytics, Tessella, part of Capgemini
Nicholas.Clarke@tessella.com

Nick leads the extensive AI, Machine Learning and Data Science community within Tessella. Nick has over 20 years' experience in leading the delivery of high-performance data-driven solutions into R&D organizations across multiple industry domains, from drug discovery through to real-time asset management in transportation systems. Nick works closely with R&D leaders to identify and implement new strategies for the innovative use of data to accelerate and de-risk the high cost, high return research programs at the heart of 21st Century Intelligent Industry.

Jerome Buvat
Global Head of Research and Head of Capgemini Research Institute
jerome.buvat@capgemini.com

Jerome is head of Capgemini Research Institute. He works closely with industry leaders and academics to help organizations understand the nature and impact of digital disruptions.

Nancy Manchanda
Senior Manager, Capgemini Research Institute
nancy.manchanda@capgemini.com

Nancy is a senior manager with Capgemini Research Institute. She keenly tracks the patterns of digital disruptions across industries to help clients understand the business impact of emerging technologies.

Ramya Krishna Puttur
Manager, Capgemini Research Institute
ramya.puttur@capgemini.com

Ramya is a manager at Capgemini Research Institute. She studies the growing role of digital technologies in transforming the boundaries of traditional business consortiums in order to help them develop informed strategies.

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/

For more information, please contact:

Global

Zhiwei Jiang
CEO, Insights and Data Global
Business Line
zhiwei.jiang@capgemini.com

Ron Tolido
CTIO, Insights & Data
ron.tolido@capgemini.com

Anne-Laure Thieullent
Vice President, Artificial Intelligence
and Analytics Group Offer Leader
annelaure.thieullent@capgemini.com

APAC

Frank Windoloski
APAC Insights & Data Leader
frank.windoloski@capgemini.com

Dr. Jing Yuan Zhao
AI Leader
jing-yuan.zhao@capgemini.com

France

Marion Gardais
AI Centre of Excellence Leader
marion.gardais@capgemini.com

India

Padmashree Shagrithaya
AI Centre of Excellence Leader
padmashree.shagrithaya@capgemini.com

Germany

Fabian Schladitz
AI Centre of Excellence Leader
fabian.schladitz@capgemini.com

Italy

Massimo Ippoliti
Vice President
massimo.ippoliti@capgemini.com

Extended Europe

Robert Engels
AI Leader
robert.engels@capgemini.com

United Kingdom

Joanne Peplow
AI Centre of Excellence Leader
joanne.peplow@capgemini.com

North America

Jerry Kurtz
NA Insights & Data Leader
jerry.kurtz@capgemini.com

Ramana Bhandaru
FS Insights & Data Leader
ramana.bhandaru@capgemini.com

Leverage Data & AI's transformative power with Capgemini

Why are Data & AI important to transform?

An entirely new breed of enterprises have emerged that are built on the activation of data, both brick-and-mortar or fundamentally digital.

These data-powered enterprises win market share:

- by delivering **better and more personalized products & services to consumers and businesses**
- by enabling the whole enterprise to leverage data to **take faster and better decisions** to drive operational excellence and competitiveness

In addition, **public sector organizations and agencies are increasingly relying on Data & AI** to automate administrative procedures, interact with citizens and state employees, detect anomalies (fraud, crime, security) and overall support decision making.

How do you succeed with Data & AI?

We believe that **All Data** (both own and external data) must be harnessed, governed and provided as **actionable insights** to an organization's business lines, for Data to be activated and AI & Analytics to be leveraged to improve on core business and to build new products and services.

Data & AI-powered transformation in order to succeed needs to be **embedded in the business**, empowered by a **data-powered leadership and culture** as well **technology usable by all, leveraging all data assets**.

But the journey is not straightforward: Artificial Intelligence (AI) implementations have now crossed the threshold of pilots, and entered the wider market. 53% of organizations have managed to scale AI projects in production – but only 13% overall have rolled out multiple AI applications across numerous teams, as per the findings of the Capgemini Research Institute report – The AI Powered Enterprise: Unlocking the potential of AI at scale.

Where to focus to deliver business impact with Data & AI?

Data is the raw material, and AI & Analytics – embedded in the business – are the set of tools to deliver tangible and trusted business outcomes.

Data activation	Data advantage	Data enablers
Data-driven end to end business solutions for transformational business outcomes	Leverage External Data, Data Partnerships & Data disruptors for competitive advantage	AI & Analytics services/insights Data & AI platforms as code

As defined in this report, Capgemini has defined 3 areas of focus that any organizations must master in order to drive Data & AI transformation forward:

Data Activation
is the construction of **insights embedded** in the organizations' line of business in order to drive operational excellence via **Intelligent Business**, **IT and Security Operations**, as well as the launch of new **Intelligent Services and Intelligent Products**

Data Advantage
is the proactive effort to leverage **external data** on top of its own data, by establishing **data-driven partnerships** with suppliers, clients or partners, including **data disruptors** such as Amazon or Google. Leveraging these external data assets in a secure and fruitful manner can be enabled via federated analytics/learning approaches and data collaboration architectures

Data Enablers
are composed of the production of **AI & Analytics insights & services** to power Data Activation, the construction of **Data & AI platforms** to support any workload at scale, and finally data strategy, management, governance, culture **foundations and behaviors** to ensure the entire organization can use Data & AI in their daily tasks.

Perform AI: Activate data. Augment intelligence. Amplify outcomes.

Perform AI, Capgemini's comprehensive portfolio of Data, Analytics, Intelligent Automation and Artificial Intelligence services, provides your organization with the ability to leverage the full transformative power of Data & AI at scale. By activating data and insights at the heart of your business, in your everyday decisions and actions, you will augment your organization's intelligence, and amplify business outcomes you expect from Data & AI.

Data & AI for trusted business outcomes

Leading organizations are today already leveraging the transformative power of Data and AI to design and launch new intelligent products, services, business models – even creating new markets. They are also finding efficiencies, optimizing processes and reducing costs, becoming more agile so they can grasp new opportunities, and developing greater resilience to sudden crises, via Intelligent Business, IT & Security operations.

Our teams have a proven track record on **Data Activation** delivering tangible business outcomes globally and at scale, across industries and sectors:

Increased sales — **25%**
Organizations successfully scaling AI have seen more than a **25%** increase in sales of their products and services

Faster customer insights — **2x** impact
A global consumer products firm successfully engaged with 1 billion people using customer data & insights with **double the impact in half the time for half the cost**.

Better customer service — **24** languages
One of the world's largest biotech companies used our award-winning* virtual assistants to automate outbound telephone calls in **24 languages**.

Operational excellence — **40%** cheaper
A leading producer of packaging and paper used our AI solution to reduce its cost of handling invoice processing queries by **40% overall**.

Cost savings — **€100m**
A major international retailer used our AI-powered solution to optimize sales forecasting which directly drove **€100m of inventory costs savings**.

Fraud detection — **10x** ROI
A European government agency has delivered a **10x return on investment** thanks to our AI-powered fraud detection systems.

Technology and Consulting excellence to empower all users with Data & AI

Combining the expertise of multiple business lines, Capgemini is able to help you get clarity on your data strategy, leveraging the technology innovations at the right pace, in a way that makes sense for your business and IT teams.

Leveraging the most powerful technology partnerships in the Data & AI ecosystems, taking into account regional specificities, Capgemini is providing you with a comprehensive and actionable framework to execute on **Data Advantage** and provide you with the right **Data Enablers**.

The right team to scale Data & AI

With **25,000 Data & AI at scale practitioners, working for more than 800 clients worldwide, supported by AI Centers of Excellence in all regions**, Capgemini's capabilities are unmatched.

We show organizations not only how transformative Data & AI could be, we actually deliver it at the heart of the business. We define the right strategy with industrialization, operationalization and scale in mind. We ensure you take the right approach to finding actionable, trusted insights for your teams, as well as including them in the transformation journey. And we build the right solutions for your business and implement them with the right continuous deployment and operations to maximize their positive impact.

Do good with Data & AI

Because we are convinced about the transformative power of Data & AI to **build positive futures for humans and society**, we partner with you to leverage Data & AI in an ethics by design and human-centered way. We work with your teams to develop the right data and AI-powered leadership, mindset, culture and ways of working that fit your organization's values. Find out more in our report AI and the Ethical Conundrum: How organizations can build ethically robust AI systems and gain trust.
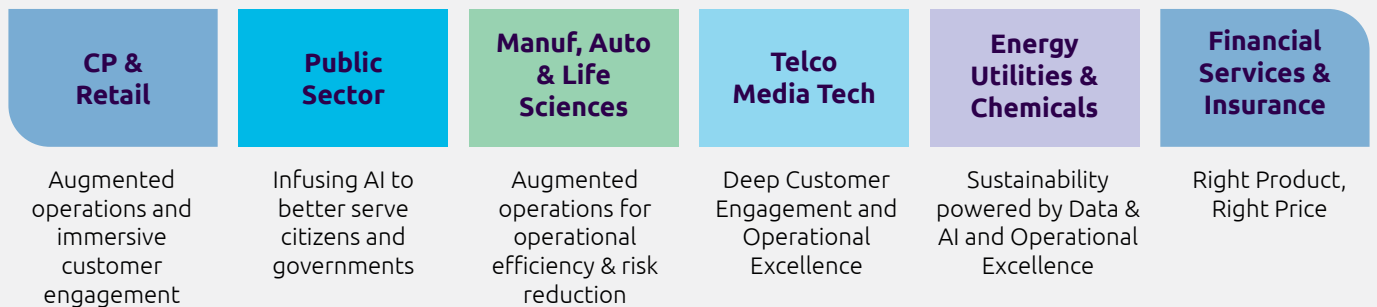
We help you balance operational excellence with business innovation to be resilient to current and future crises, to win in your market and protect your workforce. And we help you find opportunities to proactively use Data & AI for their positive impact on your business and our society.

We help you shift gears, scale trusted Data & AI solutions for the best business outcomes, and *Get The Future You Want*.
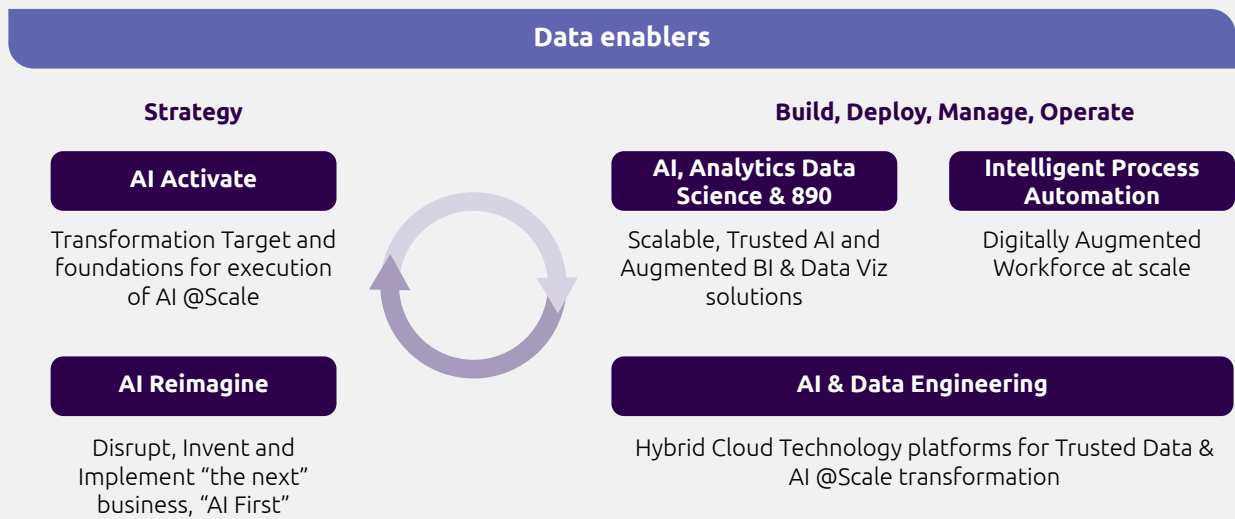
Capgemini Perform AI

Capgemini Group Portfolio framework for Data & AI powered transformation

Accelerated time to value for business outcomes with Trusted Data & AI at scale

CP & Retail	Public Sector	Manuf, Auto & Life Sciences	Telco Media Tech	Energy Utilities & Chemicals	Financial Services & Insurance
Augmented operations and immersive customer engagement	Infusing AI to better serve citizens and governments	Augmented operations for operational efficiency & risk reduction	Deep Customer Engagement and Operational Excellence	Sustainability powered by Data & AI and Operational Excellence	Right Product, Right Price

Data activation			Data advantage
Customer First	Intelligent Industry	Enterprise Management	Data Partnerships & Data disruptors

Intelligent Applications for Data & AI Powered Enterprise

Data enablers

Strategy

Build, Deploy, Manage, Operate

AI Activate
Transformation Target and foundations for execution of AI @Scale

AI, Analytics Data Science & 890
Scalable, Trusted AI and Augmented BI & Data Viz solutions

Intelligent Process Automation
Digitally Augmented Workforce at scale

AI Reimagine
Disrupt, Invent and Implement "the next" business, "AI First"

AI & Data Engineering
Hybrid Cloud Technology platforms for Trusted Data & AI @Scale transformation

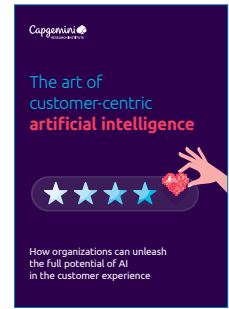| Privacy	Trust & Ethics	Security	Sustainability

Discover more about our research:

The AI-powered
enterprise Unlocking the
potential of AI at scale

AI and the Ethical Conundrum:
How organizations can build
ethically robust AI systems and
gain trust
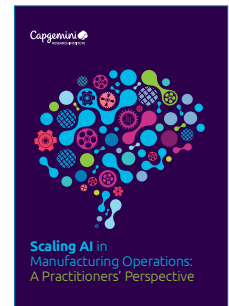
The Art of customer centric
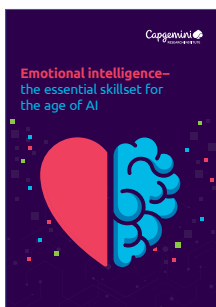Artificial Intelligence
Raising the bar on AI in
customer experience

Understanding digital
mastery today:
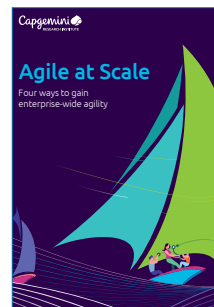Why companies are struggling
with their digital
transformations

Upskilling your people for
the age of the machine

Scaling AI in
Manufacturing Operations:
A Practitioners' Perspective

Emotional Intelligence:
The essential skillset for
the age of AI

Agile at Scale
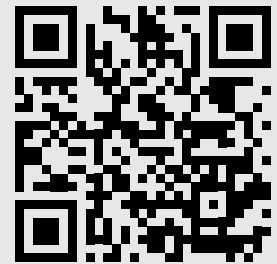Four recommendations to gain
enterprise-wide agility

The Digital Talent Gap—Are
Companies Doing Enough?

Subscribe to latest research from Capgemini Research Institute

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

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

About Capgemini

Capgemini is a global leader in consulting, digital transformation, technology, and engineering services. The Group is at the forefront of innovation to address the entire breadth of clients' opportunities in the evolving world of cloud, digital and platforms. Building on its strong 50-year heritage and deep industry-specific expertise, Capgemini enables organizations to realize their business ambitions through an array of services from strategy to operations. A responsible and multicultural company of 265,000 people in nearly 50 countries, Capgemini's purpose is to unleash human energy through technology for an inclusive and sustainable future. With Altran, the Group reported 2019 combined global revenues of €17 billion.

Visit us at

www.capgemini.com

People matter, results count.