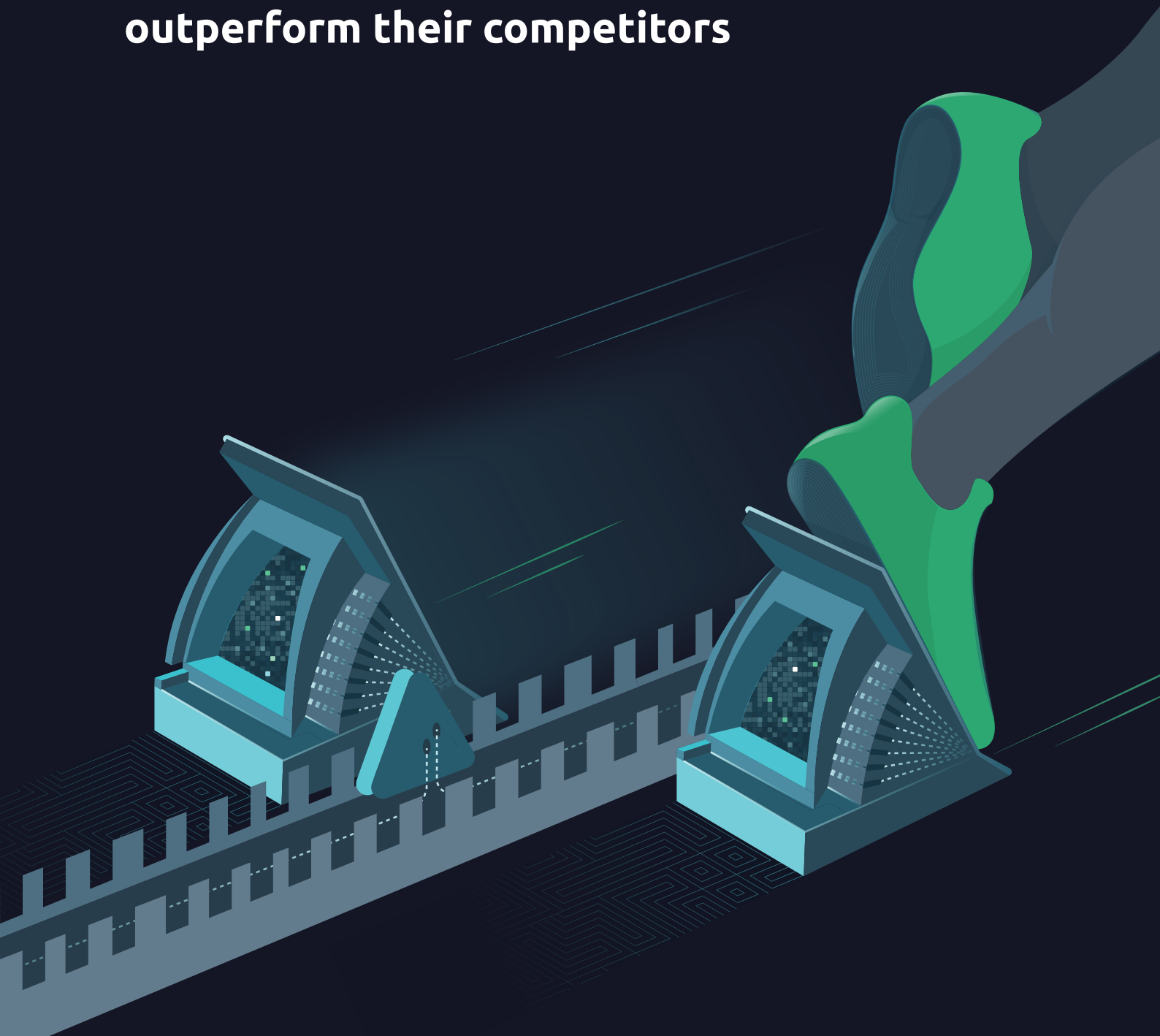


# DATA MASTERY

**How data-powered organizations  
outperform their competitors**





# INTRODUCTION

Today's business environment is both more competitive and more volatile. Traditional enterprises are constantly challenged by digital natives who have perfected the art of leveraging data-powered insights at scale, using their agility and data-powered intelligence to outperform their more ponderous competitors. At the same time, disruption, and systemic shocks – including environmental, societal, and global-health issues – are intensifying in our connected world. Organizations are forced to look towards data – along with sophisticated analytics and smart technologies such as AI – to fine-tune their strategic and tactical responses.

However, many are still far from achieving data mastery. In our previous report on **data-powered enterprises**, we assessed around 1,000 organizations across 10 sectors on their ability to derive value from growing volumes of data. We found that only a small minority (16%) could be classed as high-performers – which we call “data masters.” These were organizations that possessed the right “data foundations” – the tools and technologies to utilize data effectively – and

the right “data behaviors” – the people, processes, skills, and culture. We further found that these data masters outperformed their peers significantly on financial metrics such as profitability. For FY 2019–20, when compared to the rest, data masters on an average realized:<sup>1</sup>

- 70% higher Revenue per Employee
- 245% higher Fixed Asset Turnover
- 22% higher profitability.

In this new research, we wanted to pinpoint what sets these data masters apart. We spoke to more than 50 executives from this group:

- Technical leaders in areas ranging from information security to AI and analytics
- Functional leaders from HR to supply chain.

We also spoke to executives at big tech firms and digital natives recognized for extracting valuable insights from data.

Based on this research, we have identified four best practices:



## STRATEGIZE

### Create a data-driven strategy driven by the overall business goals

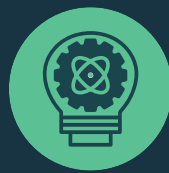
- Align your data strategy with the business strategy
- Have a forward-looking data collection strategy



## ENTRUST

### Bridge the data trust gap

- Invest in trusted data availability
- Establish guidelines for trusted AI
- Focus on data democratization



## MODERNIZE

### Revamp your data landscape

- Prioritize value streams
- Decommission legacy in phases
- Migrate to cloud and integrate multi-cloud
- Customize data-discovery tools as initiatives scale
- Accelerate innovation with DataOps



## ACTIVATE

### Drive data-powered decision making and actioning

- Create a streamlined data organization
- Facilitate data-powered innovation across the value chain
- Foster a data-powered culture across all levels of management

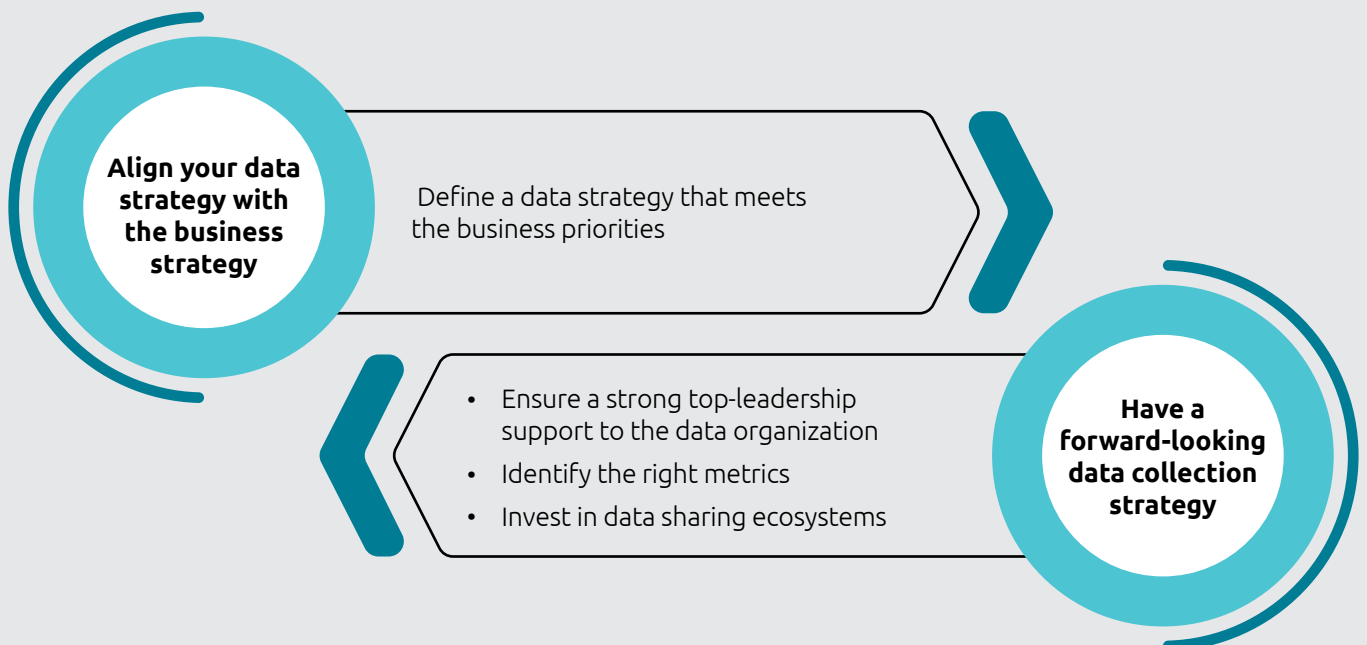
Source: Capgemini Research Institute Analysis.

In this report, we go deeper into each of the four areas.

# 1. STRATEGIZE

## CREATE A DATA STRATEGY DRIVEN BY THE OVERALL BUSINESS GOALS

A significant majority of “data masters” ensure that the wider enterprise strategy drives their data strategy. This helps them set strategic priorities and ensure foundational elements – such as measurement and collection – are therefore driving those wider goals:



Source: Capgemini Research Institute analysis.

# 1.1 Data masters align their data strategy with their business strategy

## WHY AN ALIGNED DATA STRATEGY IS IMPORTANT?

Many organizations have a data strategy. What is important, however, is its alignment with overall business objectives. IT-driven data strategies tend to focus on IT-centric outcomes, such as reducing the TCO of the data landscape or migrating to the cloud. While these outcomes are beneficial, they could lose sight of business objectives. Our previous research highlighted this issue with only 38% of business-facing executives (as opposed to IT-facing) feeling that their organization's data strategy was aligned with the overall business strategy.<sup>2</sup>



### Data masters define a data strategy that meets business priorities

Data masters start with reviewing the business's strategic priorities, from growth (generating new revenue) to operational performance (improving cost efficiency or improving sustainability). They then delve deeper into the strategic issue to define the data strategy and approach. For instance, when the objective is to enhance client acquisition by X% in a year, the data strategy could be to improve the quality of leads fed into the lead management system or to make comprehensive leads data available to sales teams.

Danone's former global Data & Analytics Officer, Ignacio Marinas outlines how Danone balances this activity, saying: "Our strategy or our mission is enabling Danone's business to take the appropriate decisions. And, based on this, we are focusing on functions such as finance, sales, and marketing and departments that deliver value across the value chain. So, the focus is then more driven by business priorities rather than by something else. From our side, the approach is also about building the foundational aspects, such as master data management, data governance, data quality, and data science."

Focusing on the specifics of the problem you are solving can ensure you do not get lost in the sheer volume of data available or become focused on the technology itself rather than the end goal. Sameer Amin, global director data driven marketing & media at Reckitt Benckiser, outlines how they used a customer-focused approach to ensure data drives specific propositions: "What you can end up doing is being infatuated by technology, rather than understanding how the technology can solve the business problem. We should instead look at what questions we want solved: What if we had X amount of data or information; what would we do with it to enhance the consumer experience?"

"For example, with Air Wick, we wanted to understand what types of scents people use in different rooms or at different parts of day. We created a connected product, and, through that, we found that people use floral or citrus scents in the mornings to revitalize, and muskier or lavender-type scents in the evening to help them sleep better," he adds.

"If you start off with technology first," Sameer suggests, "what you end up doing is buying a Ferrari and only driving in first gear."

Aligning the data strategy with the business strategy also means building data trust among the business teams. As we see in the section 2, there is a large trust gap that affects decision making.

# 1.2 Data masters have a forward-looking data collection strategy

## WHY A FORWARD-LOOKING APPROACH TO DATA COLLECTION IS IMPORTANT?



Data strategy should encompass the changing regulatory and privacy landscape as well as focus on the data elements to be collected, including ways to translate these into business metrics:

- Unless a data element can be tied to a business metric, organizations are accumulating “dark data,” and incurring costs in its storage (while increasing their CO<sub>2</sub> footprint).
- A data item earmarked for a specific purpose will be managed correctly, presented as appropriate and documented with the right business and technical metadata for clarity.

Dark data may also add regulatory compliance costs. Privacy laws surrounding consumer data require an organization to disclose to consumers how and why they are collecting certain information. Therefore, clarity on the data collected is essential.

### Data masters ensure a strong top-leadership support to the data organization

In addition to the realization that activated data is hugely beneficial for the organization, changing privacy and regulations make the ambit of data collection strategy, far reaching. There is growing support for data strategies at C-suite level, as Swapna Tom, from the enterprise data analytics team at HP, confirms: “The data-strategy team is tasked by the leadership executive team itself and they have full support from the CEO. And our strategy is forward-looking: 10 years from now, rather than three years from now. We are looking at cloud strategies, the new regulations that might emerge around the General Data Protection

Regulation [GDPR], and the data hosting itself. We are focusing on bigger, broader strategies.”

One of the principles of the GDPR states that personal data should be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed.<sup>3</sup> This not only means limiting the amount of data collected, stored, and processed, but also deleting it when it is no longer necessary. Similarly, in the United States, the California Consumer Privacy Act (CCPA) requires organizations to inform users about the categories of personal information they collect and the intended purpose. Therefore, data minimization definitely has a number of advantages:

- **Risk reduction:** Regulatory and compliance risk is minimized
- **Cost reduction:** Reduced storage and processing costs
- **Sustainability:** Reduced CO<sub>2</sub> footprint
- **Accuracy:** Better quality of data for AI and analytics use cases.

## Data masters identify the right metrics

Reporting metrics should be identified based on business requirements. Data collection and data processing should be directly tied to the business's needs, as Ram Iyer, worldwide director for digital strategy and eCommerce at Microsoft, outlines: "In my view, there are two kinds of metrics that you have to look at: performance metrics and transformational metrics. For example, I would say anything that's "sales" is really a performance metric – how are your sales trending and what's your mix? An interesting metric from the transformation perspective would be around an issue like consumer usage or renewal behavior in subscriptions."

Data masters ensure that the necessary collection, processing, and reporting processes are in place to enable data-powered decision making. When collecting and collating data from different systems, organizations must be mindful of the granularity required to take meaningful action.

## Data masters invest in data sharing ecosystems

In our connected and networked world, changes that impact suppliers, distributors, or customers will also impact the parent organization itself. Sharing data with and using data from partners boosts agility and resilience. As Figure 1 shows, our research found that a significant majority of data masters draw on a wide variety of external data sources, from publicly available competitor data to data from partners and suppliers. Our previous research on data sharing ecosystems found that organizations that use external data extensively (i.e., using more than seven external data sources) exhibit superior financial performance, i.e., up to 14x higher fixed asset turnover and 2x higher market capitalization.<sup>4</sup>

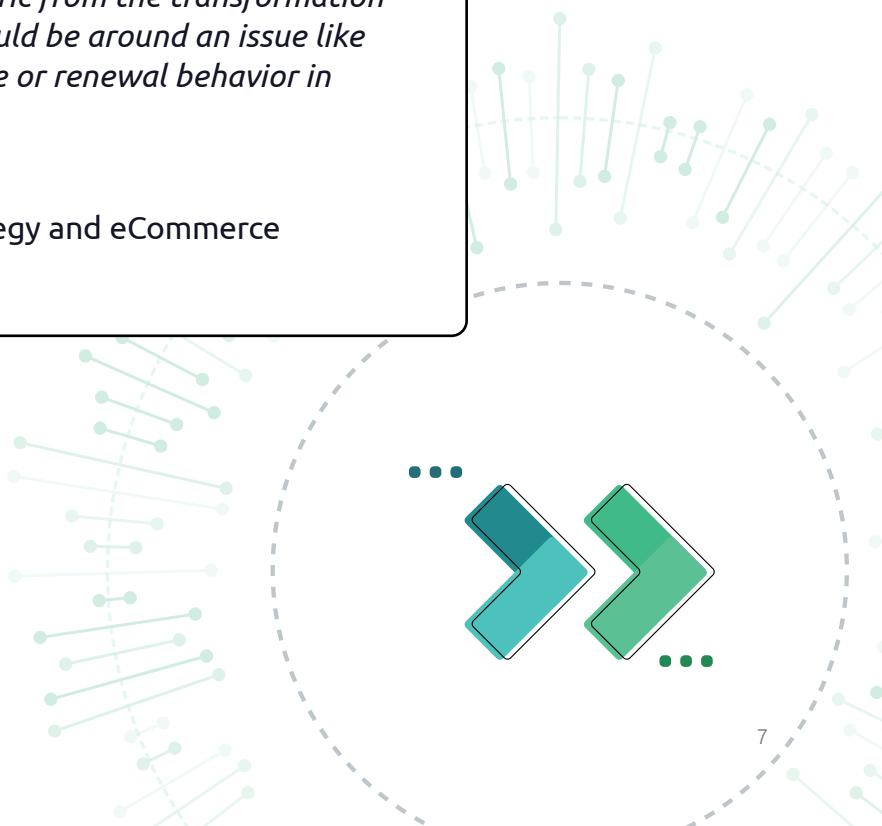
“



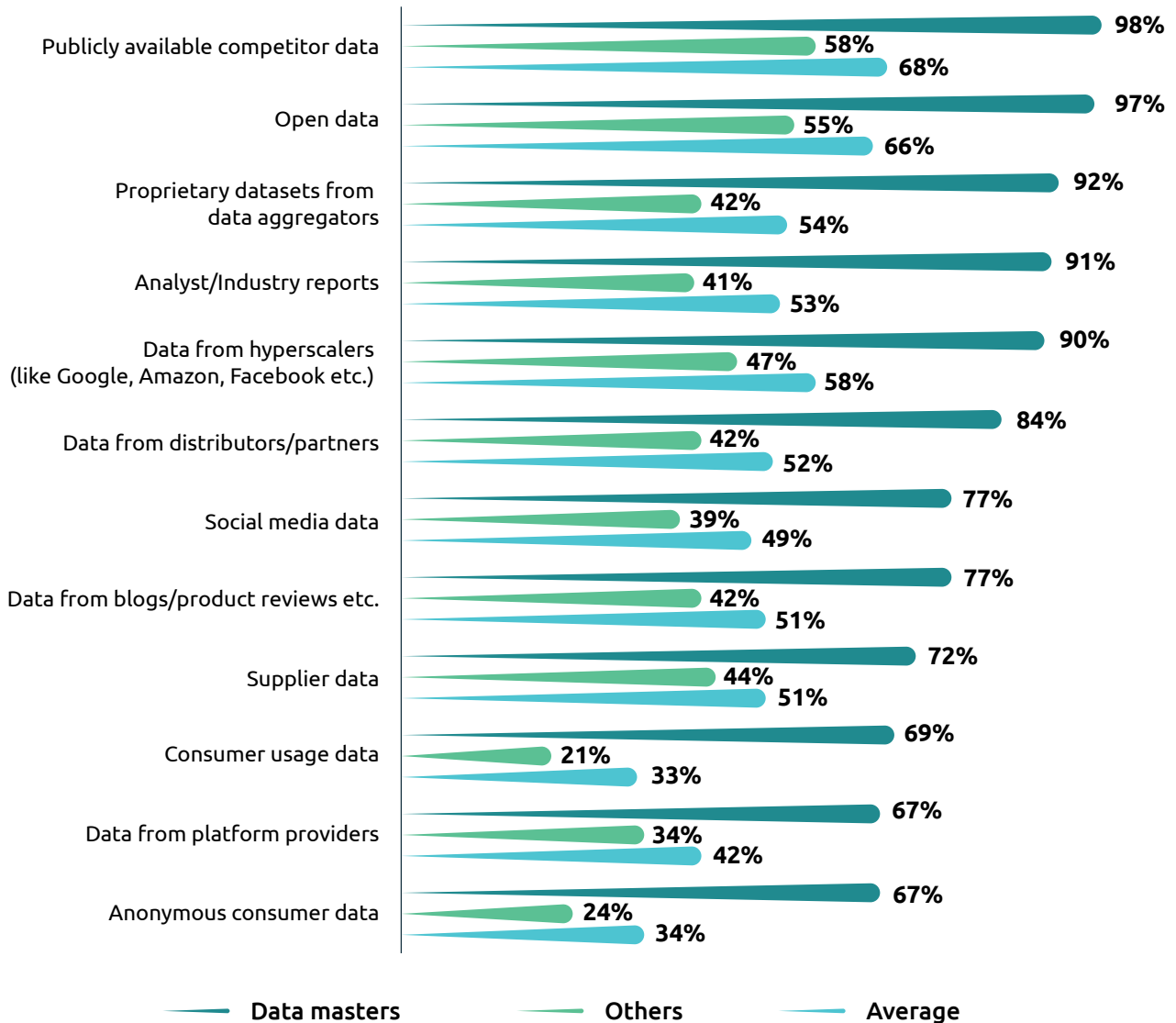
*In my view, there are two kinds of metrics that you have to look at: performance metrics and transformational metrics. For example, I would say anything that's "sales" is really a performance metric – how are your sales trending and what's your mix? An interesting metric from the transformation perspective would be around an issue like consumer usage or renewal behavior in subscriptions."*

### RAM IYER

Worldwide director for digital strategy and eCommerce  
Microsoft



**FIGURE 1** Data masters use a variety of external data sources to enhance insights



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

US-based automotive multinational, General Motors (GM), has put in place a location intelligence system that listens to a wide variety of information and generates alerts in case of adverse events. For instance, if a fire breaks out at a supplier plant and is covered by a local news channel, the team managing this system will be notified of it in near real time. To set up this system, GM worked with its tier-1 suppliers and their suppliers to collect all the location information. It also partnered with several companies to collect data from news sources and providers of critical weather information. This system improved event-detection speed and accuracy and the efficiency of GM's response.<sup>5</sup>

Reinsurer, Munich Re, developed a team called "data hunters" tasked with searching for interesting data sources (both inside and outside the company) for various use cases.<sup>6</sup>

Clearly, organizations stand to gain by investing in data sharing ecosystems. They must create a data sharing strategy that identifies the purpose of engaging in data-sharing and identify use cases that create or add value to the organization as well as to their partners. Other key areas of consideration include what data the organizations themselves should share and what the ecosystem partners would share.<sup>7</sup>





*Our strategy or our mission is enabling Danone's business to take the appropriate decisions. And, based on this, we are focusing on functions such as finance, sales, and marketing and departments that deliver value across the value chain. So, the focus is then more driven by business priorities rather than by something else..."*

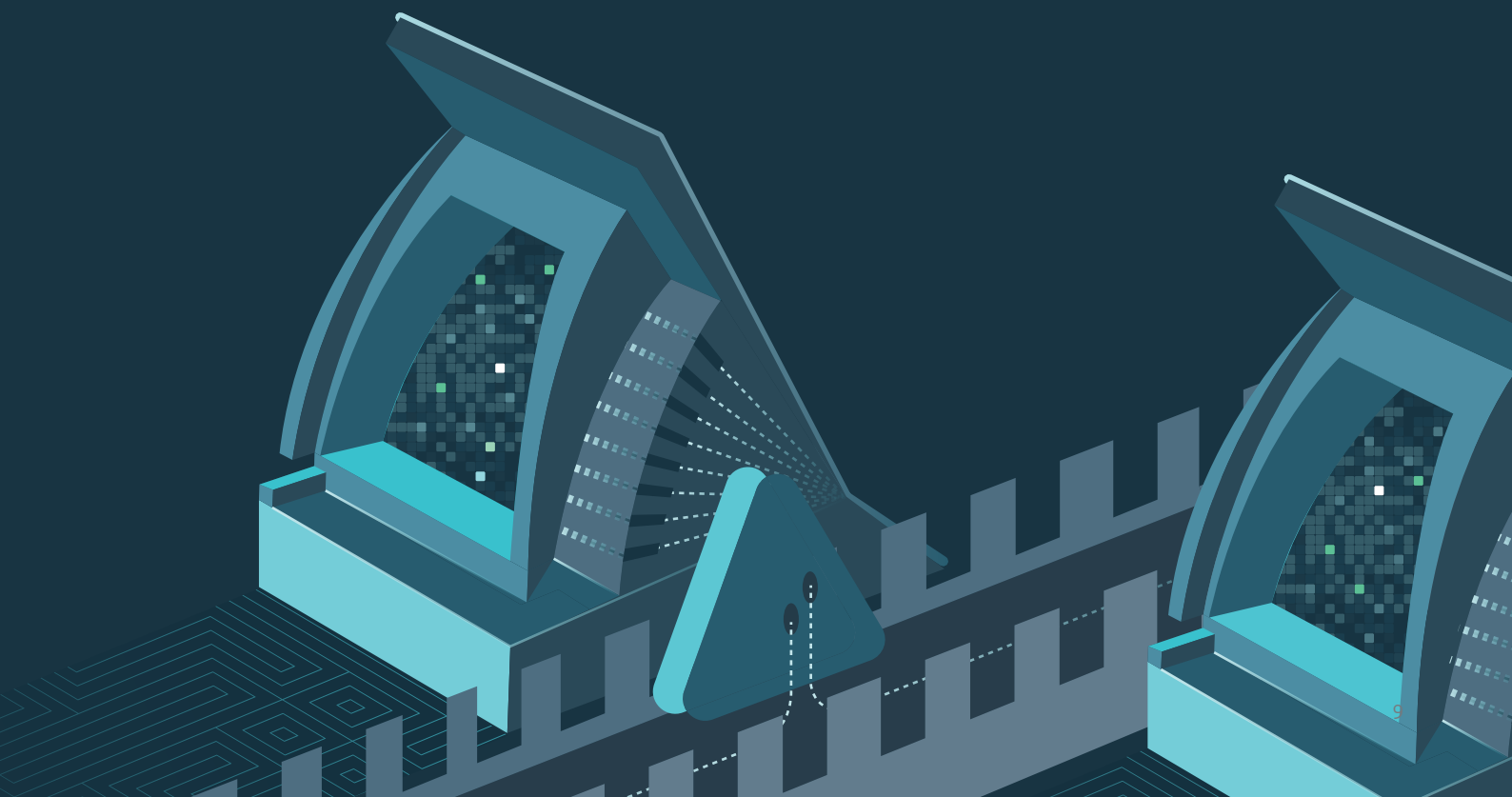
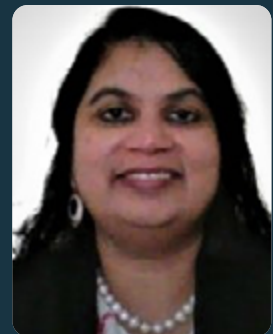
**IGNACIO MARINAS**

Former global data & analytics officer, Danone

*The data-strategy team is tasked by the leadership executive team itself and they have full support from the CEO. And our strategy is forward-looking: 10 years from now, rather than three years from now. We are looking at cloud strategies, the new regulations that might emerge around the General Data Protection Regulation [GDPR], and the data hosting itself. We are focusing on bigger, broader strategies."*

**SWAPNA TOM**

Enterprise data analytics, HP

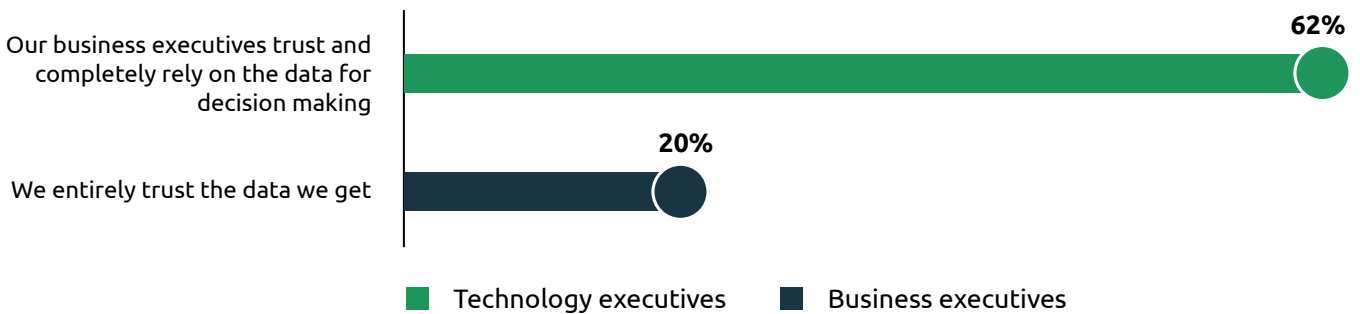


## 2. ENTRUST

# BRIDGE THE DATA TRUST GAP ACROSS AND OUTSIDE THE ENTERPRISE

Our research showed that data trust is key to organizational agility, collaboration, and generating value.<sup>8</sup> However, the research also highlighted the massive trust gap between the IT-facing arm of organizations and business units (see Figure 2). Kiran Malviya, a director at Philips, told us: “This trust gap is very evident across many organizations. As a first instinct, business wants to challenge the source of the data. They would want to challenge the accuracy, the reliability of that data. And that is why they always want to triangulate it with experiences.”

**FIGURE 2** Business executives do not trust the data



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

A managing director at a cloud provider adds: “I think the issue of trust goes back to having business at the table and driving data decisions based on business decisions, not technical ones. What tends to happen in traditional organizations is a lot of the business users, especially business leaders, see IT more as a cost center than a business

partner. IT needs to change that perception by providing value. And, it is not about just building a data store based on a bunch of technical requirements. It’s about getting value to the business by helping them become more efficient, helping them to sell more things, or helping them sell new things and create new revenue streams.”

Data masters focus on three pillars to establish trust: quality, democratization, and trusted AI:



Source: Capgemini Research Institute Analysis.

## 2.1 Data masters invest in improving trusted data availability

### WHY TRUSTED DATA IS IMPORTANT?

Trusted data acts as an accelerator to innovation. To ensure trusted data availability, organizations must focus on strong data management capabilities and invest in better data quality, given that quality is a major driver of lack of trust. The consequences of poor data quality can be enormous, with serious implications for revenue, operational efficiency, productivity, regulatory compliance, scalability and, ultimately, trust:

- **Revenue loss:** Research from Gartner has found that organizations believe poor data quality to be responsible for an average of \$15m per year in losses.<sup>9</sup>
- **Higher costs:** 20–30% of operational expenses are directly related to bad data.<sup>10</sup>
- **Productivity:** Knowledge workers waste up to 50% of their time dealing with mundane data-quality issues.<sup>11</sup>

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.<sup>12</sup>



## How are data masters improving data trust?

### Data masters focus on strong data management:

Organizations that are early on in their data-mastery journey may not have a complete picture of their data landscape, such as:

- Which data elements are digital and which are physical
- Classification of data for security purposes (confidential, internal, public, etc.) and relationships between data entities.
- Business meta-data – how data sets have been produced, when for what purposes and what use
- Technical meta-data including how data is stored and where – on premises or cloud.
- Data flow across systems/processes.

For instance, if an organization promotes its product on a social-media page, the “likes” it receives are also data that can inform analytics. Such non-traditional data sources should not be overlooked. Organizations must identify and list all the data assets they can draw on, who owns the data, and when it should be deleted in order to comply with regulation. In addition, data masters have well-defined access policies, including who can access the data and for what purpose.

Data masters build trust by focusing on:

- **Cataloging:** Data catalogs store “metadata” and help understand the various data attributes
- **Business glossary:** While data catalogs store data on data, business glossaries contain “business definitions” of the data attributes and the ‘business rules’ that can be applied to this data
- **Improving data discovery and lineage:** This includes understanding where the data comes from.

### Data masters ensure better data quality

Data masters realize that investment in data quality ultimately helps their business. Therefore, they ensure that business teams take ownership of various aspects of quality and do not see it as a completely IT-driven exercise. As Ronan Corre, enterprise data quality lead at Microsoft, told us, “Proximity with the business is fundamental because in the ocean of data, ‘boiling the ocean’ becomes a risk. So, focusing on areas most impactful for a user is key.” Data quality ranges across a spectrum of issues:

- Data-sourcing quality issues (“garbage in, garbage out”)
- Technical and business-quality issues (for instance, the definition of a “prospect” could differ in different systems)
- Data-veracity issues (accuracy, precision, and trustworthiness of data).

From our research, we found that data masters focus on the following areas to improve data quality and reliability:

#### Clearly defined nomenclature

Jean-Luc Clarot, director architecture, data and program management at Proximus, a large telco based in Belgium,

says: “Trust is also about a good alignment on the business concept itself. The root cause is not really the data itself or the way we process the data, but in the definition of the concept: what is a customer, what is your install base, from when do you understand that it’s really a new customer?”

It is critical that the business itself understands and defines what “data of good quality” is. Many data quality projects fail because the definition of data quality is monolithic and does not consider the specificities of different business teams. Take telcos as an example, where how marketing defines good data quality will differ drastically from revenue assurance:

- Marketing will be interested in analyzing overall customer use to optimize rate plans and next-best offer and will be primarily motivated by high-quality customer data.
- However, revenue assurance will be more interested in the completeness of call data records (CDRs) for all customers (they will want to analyze all CDRs regardless of quality of the customer data, as they need to detect any revenue leakage).

Data quality projects must take into account these business specificities. Otherwise, they risk being irrelevant.

#### Include quality parameters early on

Data engineers and data scientists spend a lot of time on data cleansing and preparation. However, it is important to remember that data quality means different things to different teams, so masters tend to set a clear definition early on so that it’s in place for the entire lifecycle of the data. Additionally, fixing data in the source system is often the best way to ensure effective customer experiences and analysis on the other end of the process. Swapna Tom from HP adds: “We make sure that the system of records is cleaned up first, because we believe in fixing the source, rather than the downstream.”

#### Focus on data stewardship

Designated data owners must act as custodians of data quality. Data stewards combine business knowledge and data-management knowledge, and are responsible for everything related to data collection, maintenance, organization, and compliance. Also, all users and generators of data are responsible for its quality at data master organizations such as Microsoft. “We want to create this culture that quality data is a everybody’s matter,” says Delphine Clement, director business program manager, Enterprise Data Management at Microsoft. “Everybody in the data supply chain needs to be accountable for that clean house because everybody benefits from it. For example, when we, as central data stewards, audit the quality of data, we always differentiate adherence to standards between data which has been created over the past month versus historical data. This is because we want to make sure that the recent streams of data created is better in complying to our standards than the historical data set, which is then a testimony that our initiatives for establishing a data quality accountability culture work.”



## LEARNING FROM DATA MASTERS: USE OF AI FOR DATA HYGIENE

Fixing the data-integrity issues at point of collection (or soon after) helps to ensure data hygiene and provide quality data for business intelligence. However, maintaining data integrity is often a challenge when dealing with scores of external siloed data sources. Issues such as transfer error, human error, or misconfigurations can make the integrity-checking process arduous. Organizations such as Microsoft are turning to AI to help.

### Use AI to check data integrity

“Our sales data, from our eCommerce partners worldwide, comes into data engine MS Sales which is the single repository” says Microsoft’s Ram Iyer. “Once the data comes in, it is important to do integrity checks. This is where we can use the power of AI. Now, you have really good AI tools that can scrub data and potentially measure and improve the quality of data with a data-hygiene index.”

## 2.2 Data masters establish guidelines for trusted AI

### WHY ETHICAL AI IS IMPORTANT?

As we have seen, lack of data trust is hindering businesses in capitalizing on data insights. While the quality of data is one factor that impacts data trust, other issues include the “black-box” nature of AI algorithms and possible inherent bias. Unless business executives have a clear understanding of how AI works – and are confident in the fairness of the AI systems – they may not be able to trust the output of these algorithms.

Furthermore, the past few years have seen growing concern around the ethical aspects of AI. By analyzing 500,000 blogs and 60,000 English news stories, Stanford University’s 2021 AI Index report found that articles related to AI ethics stories topped the list of most-covered AI topics in 2020.<sup>13</sup> This showcases the growing interest in ethical AI.

For more details on how we define ethical AI please see our [Code of Ethics for AI](#) (see also Appendix 2).

In addition to executive trust, another critical aspect of trusted AI is the growing awareness of ethical concerns among consumers in their use of AI systems. For instance, our research on “ethics in AI” showed that 71% of consumers want an organization’s AI systems to be able to clearly explain the results it arrived at, and 67% expect organizations to take ownership of their AI algorithms when they go wrong.<sup>14</sup>



## How do data masters ensure their AI is trusted?

### Data masters set up robust processes for governing AI

Ensuring AI solutions are trusted involves a series of actions. Some of the best practices that data masters follow include:

#### Define an AI charter

Organizations must determine what AI ethics means to them, and how they can define and enforce it. While regulation provides guidance, it is up to organizations to define their AI codes of conduct. Nicolas Economou, co-chair, Law Committee of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, and CEO of H5, explains: “Beyond what you must comply with, you need to determine what you stand for as an organization. What are your brand values? You can then define your ethical AI code on the basis of what you stand for; what that implies for how you think of the impact of your decisions on your company, employees, customers, and society at large.”<sup>15</sup>

Mohsen Rezayat, chief solutions architect at Siemens Digital Industries Software, says: “Ethical considerations need to be thought about at every stage of product development. We are investing in research to enhance robustness and security, explainability, safety, privacy, and accountability (including guarding against harmful bias and unintended consequences) to make AI solutions ethical and trustworthy in a B2B context.”<sup>16</sup>

#### Establish leadership and governance

Data masters establish a position that oversees the data and AI ethics. For example, H&M, the Swedish clothing-retail company, appointed a head of AI policy and created a checklist that is used for all ongoing and new AI projects. This checklist is centered around nine areas: focused, beneficial, fair, transparent, governed, collaborative, reliable, respecting privacy, and secure.<sup>17</sup> Such a position helps define the AI charter and drive the ethical use of AI across the organization.

Data masters also establish mechanisms to raise and address the ethical concerns around the development of AI. These

could mean creating hotlines or channels for employees to speak about various ethical issues they may face during the design, development and implementation of AI systems.

#### Set up processes around AI design and development

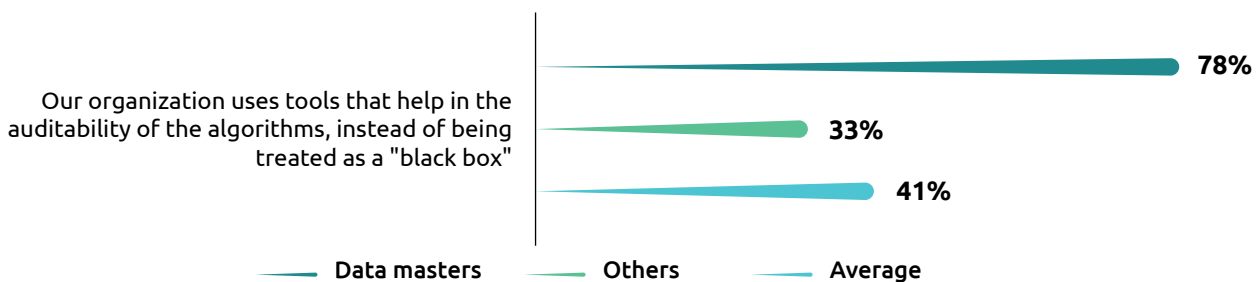
Having a set of guidelines and processes around AI design and development will be beneficial for the organization in the long term. Such practices include:

- Laying out the purpose of an AI system and communicating it transparently to all the stakeholders
- Identifying and preventing accidental or malicious misuse: In 2014, Amazon started working on AI-based program for recruitment to review the resumes of applicants and select the top profiles. However, the AI results were gender-biased and after failing to guarantee that the program would remain non-discriminatory, the AI system was disbanded altogether.<sup>18</sup>
- Ensuring accountability of AI systems – human oversight through mechanisms such as human in the loop (HITL), human on the loop (HOTL) or human in command (HIC) helps in maintaining control and accountability.
- Building diverse and inclusive teams – Lack of diversity in AI teams often causes any inherent bias to remain hidden. By bringing in diversity into the design and the development teams, both in terms of people as well as in discipline, organizations can make fair AI possible.
- Auditing AI systems for adherence to the AI charter or the code of conduct.
- Training programs for sensitizing both developers and management on various ethical aspects of AI. Google, for instance, trained 5,000 customer-facing cloud teams in asking critical questions to spot potential ethical issues.<sup>19</sup>

#### Use tools and frameworks to develop ethical AI systems

While an AI charter helps guide the teams, equipping the development teams with tools and frameworks helps them implement the concepts of ethical AI. For instance, our research shows that data masters rely on tools and frameworks that help them in auditability of AI. As Figure 3 shows, four in five data masters rely on such tools, compared to just 33% of others.

**FIGURE 3** A majority of data masters rely on tools to help the auditability of AI algorithms



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

Explainable AI-as-a-Service is offered by Amazon, Google, and Microsoft among others. Google's Explainable AI, a set of tools and frameworks to help analysts understand and interpret predictions made by their machine-learning (ML) mode, includes features such as:<sup>20</sup>

- A score explaining how much each factor contributes to the model predictions
- A "what-if" tool to help investigate model performances for a range of features in the dataset, optimization strategies, and even manipulation of individual datapoint values.

## 2.3 Data masters focus on data democratization

"Democratization" of data is the ability to provide the required data 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. This not only empowers the business but also makes them more accountable for driving data-powered decision making.

### WHY DEMOCRATIZATION OF DATA IS IMPORTANT?

- The ability to access and understand data translates into **faster decision making and greater agility**. German pharmaceutical company, Boehringer Ingelheim, has democratized data by sharing metadata across clinical-trial phases to provide real-time data to researchers. It has since achieved a better and faster data flow in its clinical-trial processes and accelerated its drug development pipeline.<sup>21</sup>
- Arming customer-facing departments, such as customer services and sales, with accurate customer data allows them to deliver a **seamless customer experience**.
- Democratization creates more **empowered, engaged, and accountable employees**. Bernard Marr, author of *Big Data in Practice* (2016), says: "**When you allow data access to any tier of your company, it empowers individuals at all levels of ownership and responsibility to use the data in their decision making.**"<sup>22</sup>
- And finally, it is **key to establish data trust**. Himanshu Shekhar, vice president, enterprise data, analytics and insights at PVH, a US apparel company that owns brands such as Van Heusen, Tommy Hilfiger, and Calvin Klein, comments: "**Trust has a lot to do with data transparency – with data democratization. The lack of trust from senior leaders comes from the fact that they are reliant on individuals to provide the information to them, and then different individuals start providing different pieces of the puzzle, with different filters or in a different interpretation. This leads to executives lacking trust in the data or a desire to extrapolate results themselves.**"



## How are data masters democratizing data?

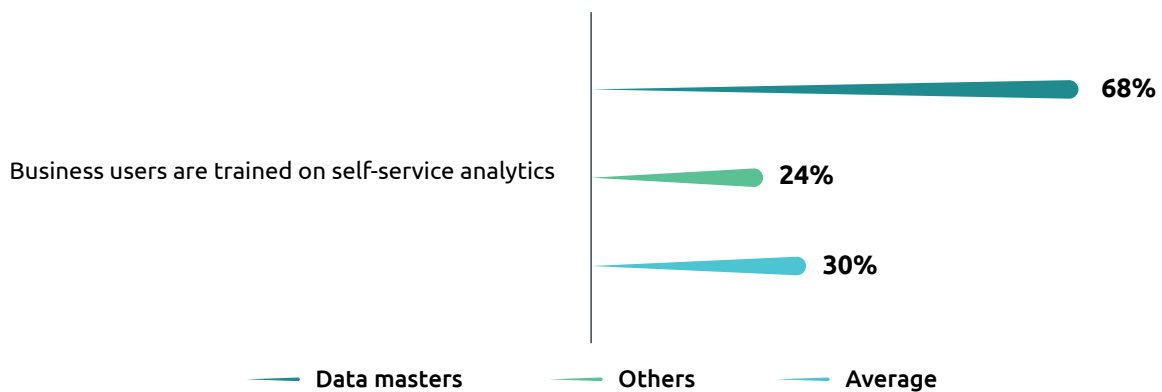
### Data masters empower businesses through self-service:

Through the availability of trusted data, data masters strive to make the business teams agile. Andreas Kohlmaier, head of data engineering at Munich Re, the multinational insurance company, outlines an example of how their group-wide data platform promotes collaboration and innovation, saying: “One of the things that happened recently is a use case about offering cover for wind farms. On one side, we have our asset-management team that actually owns and runs

wind farms. Then, we also have a subsidiary in Hartford in the US that specializes in IoT and in monitoring wind farms for other customers, and we have a weather team in Houston that is expert on wind and climate. Those three groups came together and combined their knowledge about their experience in running wind farms, IoT, the sensor data, and the weather. They combined it in a new service product for anyone who plans to invest, run, and build a wind farm.”<sup>23</sup>

As figure 4 shows, data masters invest in training business users in self-service analytics.

**FIGURE 4** Data masters train their business users in self-service analytics



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

Some best practices emerged from our discussions:

- Do not consider self-service analytics as a one-time project:**  
 GE-Aviation, a provider of jet and turboprop engines, uses two permanent teams that support their robust self-service initiative:
  - The Self-Service Data Team, responsible for user enablement, tool administration, and identifying opportunities for process automation
  - The Database Admin Team, responsible for data governance and user support.<sup>24</sup>
- Clearly define a glossary of terms for business users:**  
 Kurt Muehmel, chief customer officer for Dataiku – an artificial intelligence and machine-learning company providing platform for democratizing access to data – told us: “In many organizations, generally among business users, there is a problem of understanding the context – where the data is coming from, why the data is structured the way it is, and what would be appropriate uses of that data. So, I think the solution here is ensuring that alongside the data itself, organizations have embedded the context and the understanding necessary for how to use it.”

- Focus on education as a means to enable self-service:**  
 Data masters who embark on a journey to provide self-service data to business users do not consider intuitive tools a replacement for training and education around the initiative. The self-service data (SSD) program in GE-Aviation focuses on education to obtain high adoption among business users, as well as sustained use of the tools and program over time:
  - Onboarding: Regardless of role and background, all GE SSD users go through the Digital Data Analyst (DDA) training as a part of onboarding
  - Ongoing: In addition to DDA, the SSD team also spends a lot of time on improving materials on ongoing education and training
  - Gamification: The team launched a “Data Duel” challenge – where users were given points for activities such as tagging a dataset, creation of new dataset or documentation, etc. More points unlocked the possibility to pass levels and get exclusive laptop stickers. Leaderboards were also used to keep the interest and engagement high.<sup>25</sup>



**Data masters focus on data visualization and storytelling:**

Data-visualization solutions are the tools used to convert raw data into visuals that are easy to interpret by the business – consequently, they are critical to data-democratization efforts. They can be called the “face” of self-service analytics. Data masters employ these visualization tools to make data-powered insights understandable and actionable to everyone.

**Data masters improve data citizenship through reskilling and training:**

We define “data citizenship” as understanding the inherent potential of data, interpreting it accurately, and driving decision-making and business users with this agency as “data citizens.” As mentioned above, self-service analytics and

visualization tools have been key to data masters’ fostering data citizenships, so that business users can create a narrative based on the data. Data masters emphasize the importance of widespread data literacy to build this “data-first” culture – as 87% of data masters expand their data-literacy programs beyond the subject-matter experts. As Aleksejs Plonitkovs, enterprise data director at Microsoft, explains: *“We put a significant effort into educating our business communities and our business users about the importance of data, governance principles, and the importance of having clean datasets. We call this as ‘efforts towards data literacy.’ We want to be sure that every Microsoft employee is aware of the principles of data governance and data management and how to sustain high quality of the data.”*

“



*We put a significant effort into educating our business communities and our business users about the importance of data, governance principles, and the importance of having clean datasets. We call this as ‘efforts towards data literacy.’.....”*

**ALEKSEJS PLONITKOVS**  
Enterprise data director,  
Microsoft

“

*In many organizations, generally among business users, there is a problem of understanding the context – where the data is coming from, why the data is structured the way it is, and what would be appropriate uses of that data. So, I think the solution here is ensuring that alongside the data itself, organizations have embedded the context and the understanding necessary for how to use it.”*

**KURT MUEHMEL**  
Chief customer officer,  
Dataiku



# 3. MODERNIZE REVAMP YOUR DATA LANDSCAPE

## 3.1 Data masters modernize their data estate

### WHY REVIEWING THE DATA ESTATE IS IMPORTANT?

In an omnichannel world, customers expect to move seamlessly across offline and online touchpoints. A modern technology platform boosts the organization's ability to respond to shifting customer behavior and generate actionable, timely insights. It includes the ability to combine data from multiple internal and external sources and deliver insights at speed and scale.

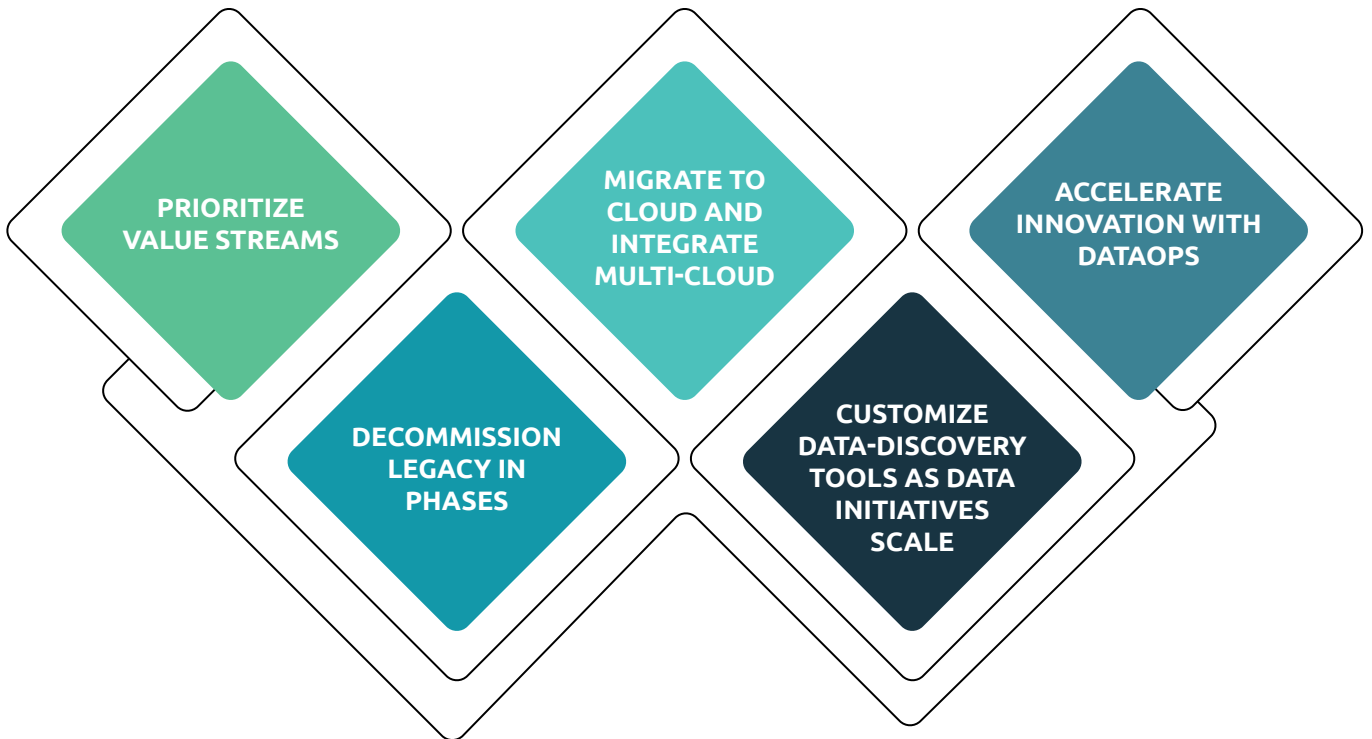
CarMax, an automotive retailer, has adopted many leading practices in software, including continuous integration, continuous delivery, application programming interfaces (APIs), and microservices-based architecture. It enabled the organization to improve code-release frequency from once every two months to multiple times a day. This constant improvement allows the company to roll out products, updates, enhancements, and fixes to market much more quickly, improving customer experience.<sup>26</sup>



### How are data masters modernizing their data estate?

Most companies have already accepted the need for digitally enabled IT systems. However, the required multi-year transformation often costs tens of millions of dollars and leads to loss of competitive position. Also, data-platform tools and technologies evolve rapidly, and companies need to make crucial build-versus-buy decisions to keep their investments aligned with the technologies they need.

Data masters focus on the following areas to deliver business priorities:



**Source:** Capgemini Research Institute Analysis.

**Data masters prioritize value streams (products, services, or a portfolio of these)**

As we saw in section 1, data masters ensure alignment between the business strategy and the data strategy. The business objectives can be aligned with data priorities by mapping value streams over the modernization efforts. To achieve this, they focus on the following elements:

- In data masters, these value streams emerge from larger business priorities, such as expanding market share, reducing platform costs, improving customer engagement, and automating processes. They also help in early value realization.
- Data masters logically group applications and infrastructure required to launch applications, and provision datasets required for prioritized value streams.

**Identify and prioritize value streams that are most impactful for the business**

- At an early stage, data masters remove legacy dependencies and abstract applications using APIs for delivering benefits.
- They also identify and fill gaps in capabilities necessary to deliver those value streams.

**Use value streams to pinpoint system dependencies**

Prioritization of these value streams helps bring the business and technical teams onto the same page, and governance structures should therefore include representatives from both sides. The overall estate modernization roadmap should be linked to the delivery of such value streams as milestones to ensure strong alignment with business objectives. Further, data masters ensure that they have a backlog of these value streams with defined priorities.

### Data masters decommission legacy systems in phases

Companies that have worked with monolithic systems for years cannot practically uproot these and replace them with modern platforms. Companies start out by mapping relevant transactional and historical data with the new systems and

transforming data for use in the new systems. After this, they divide the monolith into containers of larger components, which can be used independently.

Canadian eCommerce retailer, Shopify, has adopted a two-pronged approach to decoupling its legacy systems. It started out by drawing boundaries through the monolith, consisting of related components. Gradually, it strengthened those lines to create full-fledged boundaries with a stewardship team assigned for each component. Secondly, it carved out isolated and clean components from its legacy codebase, where relevant.<sup>27</sup>



## LEARNING FROM DATA MASTERS: DECOUPLING WORKSTREAMS FROM LEGACY IT

DoorDash, a US-based online food ordering and delivery platform, shifted its pricing algorithms out of the monolith to quickly test new pricing components.<sup>28</sup>

### Challenge

DoorDash's pricing algorithm was implemented in a legacy monolithic codebase, which caused a number of issues:

1. When customers checked out with multiple items in their cart, each line item had duplicated implementations dispersed throughout the codebase.
2. Legacy code became unwieldy and difficult to understand over time, owing to technical debt (rising complexity of algorithms over time) and multiple engineers working on it.
3. The legacy systems were not sufficiently scalable to accommodate the increase in traffic.
4. Owing to tightly integrated modules, platform-wide consistency was missing when calculating prices. Introducing changes in pricing algorithms was complex, as it involved altering tightly coupled components across a large codebase.

### Solution

To address these issues, DoorDash separated the pricing logic out of the monolith using microservices architecture. Creating a pricing service as a platform led to a centralized and standardized framework that allows engineers to implement and test their new pricing components and use cases clearly and rapidly.

## Data masters manage data as a strategic asset across multi-cloud environments

Cloud migration is already a priority at many large organizations. Appetite for cloud-based systems is driven by a range of advantages:

- They are not only cost-efficient but help drive innovation.
- They enhance analytics capabilities at scale; increase development velocity; and harmonize datasets across multiple environments to provide a consistent view of business.

Seeing these benefits, US-based CapitalOne Financial, for example, recently moved 100% of its data to the public cloud.<sup>29</sup>

Similarly, US-based insurance group, Nationwide, moved over 850 business and customer-facing applications – such as claims, personal, and commercial insurance policy systems – to cloud. With this, it could deliver deeper insights to its independent agents, financial professionals, and customer-service associates to provide tailored recommendations to small-business customers within minutes.<sup>30</sup>

Some leading practices emerge from our research with data masters:

**Create an integrated data management across the organization:** Data masters realize that to achieve scalability and agility, data ownership should be federated, with each business domain handling their own data sets. However, they also allow for interoperability across these different domains by applying common data standards. Data mesh architecture, for instance, allows federated data ownership while allowing communication between distributed data sets.<sup>31</sup>

- Cloud migration helps in improving agility of data access. For example, Ricoh, an imaging and electronics company, implemented cloud to connect all data from its factory to deliver process improvement, enhance product performance and lower the failure rate of machines. Telemetry data on the performance of machines is ingested in real time in cloud and live adjustments are made to ensure optimum performance and overall product quality.<sup>32</sup> As Ignacio, formerly from Danone, points out: “We are trying to get away from on-premises storage tools and bring all the data lakes to the cloud. With that, we have access from everywhere and can ensure security of data. Through our mapping tool, everyone with permissions is able to access data from anywhere.”

- Large organizations can also have multiple data stores that cater for the unique needs of different product groups or functions. These data stores also connect to the enterprise-wide data store. As Aleksejs from Microsoft told us: “Our enterprise data lake comprises a lot of different sources, albeit not everything. There are multiple data lakes because of the size and of specialty. For example, we have dedicated financial data lake which is more focused on the financial data, as well as own repository of data rules, governance principles, metadata and taxonomy definitions for the company.” Ensuring that the definitions are standard across these data stores is critical so as to avoid misinterpretation of the data.

### Plan for interoperability within multi-cloud environments:

Companies usually select a keystone cloud vendor to migrate their systems, code, and applications to public or private cloud environments. They also keep peripheral vendors to get the best of capabilities from each vendor, reduce concentration risk, and meet regulatory requirements.

Data masters ensure that data residing with multiple cloud vendors behaves as if belonging to one source. Interoperability can be improved by working with open-source or package solutions using the same codebase and APIs to enhance integration among cloud vendors. Open APIs, for instance, use a consistent platform to deploy and manage cloud-native and legacy applications across multiple clouds with a single interface. Furthermore, data teams can use established and verified codes from libraries, instead of building from ground zero, which quickens development velocity and reduces errors.

### Use “federated learning” to work with distributed datasets:

Transferring high volumes of internal and external data is costly, difficult, and often risky from a security, privacy or compliance perspective. Federated learning overcomes this; it is an approach that makes it possible to train ML models on distributed data while preserving the privacy of data subjects.

For example, data from different hospitals cannot be easily combined owing to a plethora of privacy rules and regulations, but it is critical to improving patient outcomes. Using federated learning, hospitals are now able to train their algorithms on data from multiple institutions. At the same time, they can keep their data and proprietary algorithms confidential. Intel launched a federated-learning model using patient data from 23 hospitals, which was combined to improve accuracy of anomaly detection in brain scans by 11%.<sup>33</sup>



# LEARNING FROM DATA MASTERS: DESIGNING FOR INDUSTRIALIZATION THROUGH EXPERIMENTATION GUARDRAILS

Data masters understand that scaling a data culture is about creating policies and processes for fast experimentation. They achieve this by establishing “guardrails” – policies and standards for rolling out new features/products/services to market. Data teams spread across the organization can design and run experiments within the parameters of these guardrails. This ensures velocity is maintained but risk management is not forgotten. Airbnb shows the power of this sort of approach in action:

Teams across Airbnb run thousands of experiments on its website every week. Each experiment, aimed at improving a particular metric, can affect key metrics of other teams. For example, the trust team may prioritize fraud identification, while the experiences team may prioritize discovery of the online experience on the company’s homepage.

To tackle this situation, in 2019 Airbnb launched an Experiment Guardrails system:

- The system has three sub-guardrails, and each experiment has to pass through all the filters. One of these is the Impact Guardrail, which requires that the change anticipated by a particular experiment does not affect other key metrics above a specific percentage.
- If a team wishes to launch an experiment that has “triggered” a guardrail (where the guardrails system finds that it negatively impacts a key metric), they will initiate an escalation process, where stakeholders can discuss the results transparently.
- The system flags roughly 25 experiments per month for escalation or review. And, of these, 80% eventually launch after discussion between stakeholders and additional analysis, and 20% (five per month) are stopped before launch.

This configurable system allows Airbnb to balance two priorities: safeguarding key metrics while maintaining speed of product development.

**Sources:** InfoQ, “Airbnb: Using Guardrails to Identify Changes with Negative Impact across Teams,” January 2021. News articles.

## Data masters customize data-discovery tools as data initiatives scale

Data discovery implies collecting, evaluating, and recognizing patterns in data from various sources. Organizations use it as a framework through which to understand their data. The data-discovery process includes connecting multiple data sources, cleansing and preparing the data, sharing the data throughout the organization, and performing analysis to gain insights into business processes.

Life sciences major Bayer has automated its data discovery process as Holger Buchner – Pharma IT enterprise architect at Bayer – has stated: “The data transport is now automated, which reduces the amount of discovery work and research for the interface development to a minimum and lets people focus on the core of the business: understanding how business data fits to each other between applications.”<sup>34</sup>

Given the critical nature of this operation to understanding data and deriving insights, digital-native companies such as Facebook, Airbnb, Lyft, Netflix, LinkedIn, Twitter, Uber, and Shopify have either customized or built their own solutions for data discovery.<sup>35</sup>

Before building data-discovery platforms, data teams can struggle to understand whether there is an existing data asset that can be utilized for the current problem, or which teams will be affected by changes as a result of their implementation. At Shopify, this involved asking team members and sifting through internal code logs, resulting in multiple sources of truth, lack of full context, and duplication of effort. When Shopify created its data-discovery platform, it increased productivity of data teams, reduced business owners’ dependence on data teams, and improved accessibility of data.<sup>36</sup>

### Data masters accelerate innovation with DataOps

DataOps requires a fusion of technology, processes, and people. We define it as: “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.”

As data masters seek to become ultra-agile, ultra-adaptive, and ultra-responsive, they look towards tightly integrated, multi-disciplinary teams that rapidly bring solutions to operations. They are employing DataOps to connect data consumers with data creators to enable collaboration and accelerate innovation based.

DataOps are helping data masters to achieve a number of aims:

- Provide faster access to actionable intelligence: 85% of data masters are deploying DataOps practices to improve the quality and speed of end-to-end data pipelines and 90% are using it to deliver analytical solutions quickly
- Ensure continuous data flows by automatically identifying and handling data drift. ‘Data drift’ refers to the change in model input data that leads to model performance issues. Ericsson’s director of data science, Dr. Sunil Kumar Vuppala, says: “A couple of years ago, concept drift and data drift were not our priorities. But now, we are advocating that the data is going to change over time, and COVID is the best example of patterns changing and of concept/data drift. Here, MLOps and DataOps are really helping us – as we talk about real model deployments and democratization.”

Some best practices emerge from data masters for a successful enterprise-wide DataOps strategy:

- **Build a culture of collaboration with cross-functional teams:**

DataOps must drive collaboration with the business-unit stakeholders. One large financial institution has built a data lab that creates products and solutions that harness data and advanced analytics to allow a deeper understanding of its 60m+ customer base. The data lab has taken a holistic approach to achieving this goal, starting with mingling disciplines such as human-centered design, full-stack engineering, and data science, and then working continuously to build an interdisciplinary team. The team includes a project manager who oversees the entire end-to-end pipeline, as well as several DevOps team members and data scientists.<sup>37</sup>

- **Create a DataOps unit:**

Data masters create a DataOps department headed by a CDO. MoneySuperMarket, a British price-comparison website, has created a data team outside of IT to accelerate the company’s data-science initiatives. The reorganization pulled in data specialists from IT and other areas to create a team dedicated to data science.<sup>38</sup>

- **Focus on metadata management for a mature DataOps capability:**

A senior data executive for a large Asia-based bank told us: “You need to define data. First the metadata should be defined and, ideally, there should be a platform within the data platform doing metadata management.”

- **Automate provisioning of data, analytics, and AI infrastructure:**

To support the integrated-care model, Humana, a US-based healthcare company, created a hyper-connected ecosystem and applied DataOps principles and tools. It automated and accelerated the provisioning of data from multiple sources, in a variety of formats, to deliver it into a hybrid-cloud environment.<sup>39</sup>



*A couple of years ago, concept drift and data drift were not our priorities. But now, we are advocating that the data is going to change over time, and COVID is the best example of patterns changing and of concept/data drift. Here, MLOps and DataOps are really helping us – as we talk about real model deployments and democratization.”*

**DR. SUNIL KUMAR VUPPALA**

Director of data science,  
Ericsson

# 4. ACTIVATE

## DRIVE DATA-POWERED DECISION MAKING AND ACTION

As organizations incorporate data from third parties and work to build their own repositories, they gain knowledge – about competitors, customers, and the world in which they operate. At the same time, simply accumulating data is not enough. Organizations need to “activate” this data.

“Activated data” is about transforming data and insights into actions within an organization’s processes and ways of working that helps them to predict market dynamics, anticipate trends, identify customer behaviors, manage risk, find operational efficiencies, grow, and innovate.

Activating data for “optimization” – for improving process efficiencies, business performance, and profitability – is not a new territory for data masters. Our research highlights that, with activated data, data masters on average see a 16% increase in operational efficiency. For instance, by activating data gathered from internet of things (IoT) sensors in parcel-processing facilities and delivery vehicles, UPS has been able to reduce average delivery time by a day in 63% of the zip codes it serves.<sup>40</sup> While activation of data for decision making and monitoring performance will always be key, data masters are now looking at transforming the data into new revenue streams and new business models (see Figure 5).

“

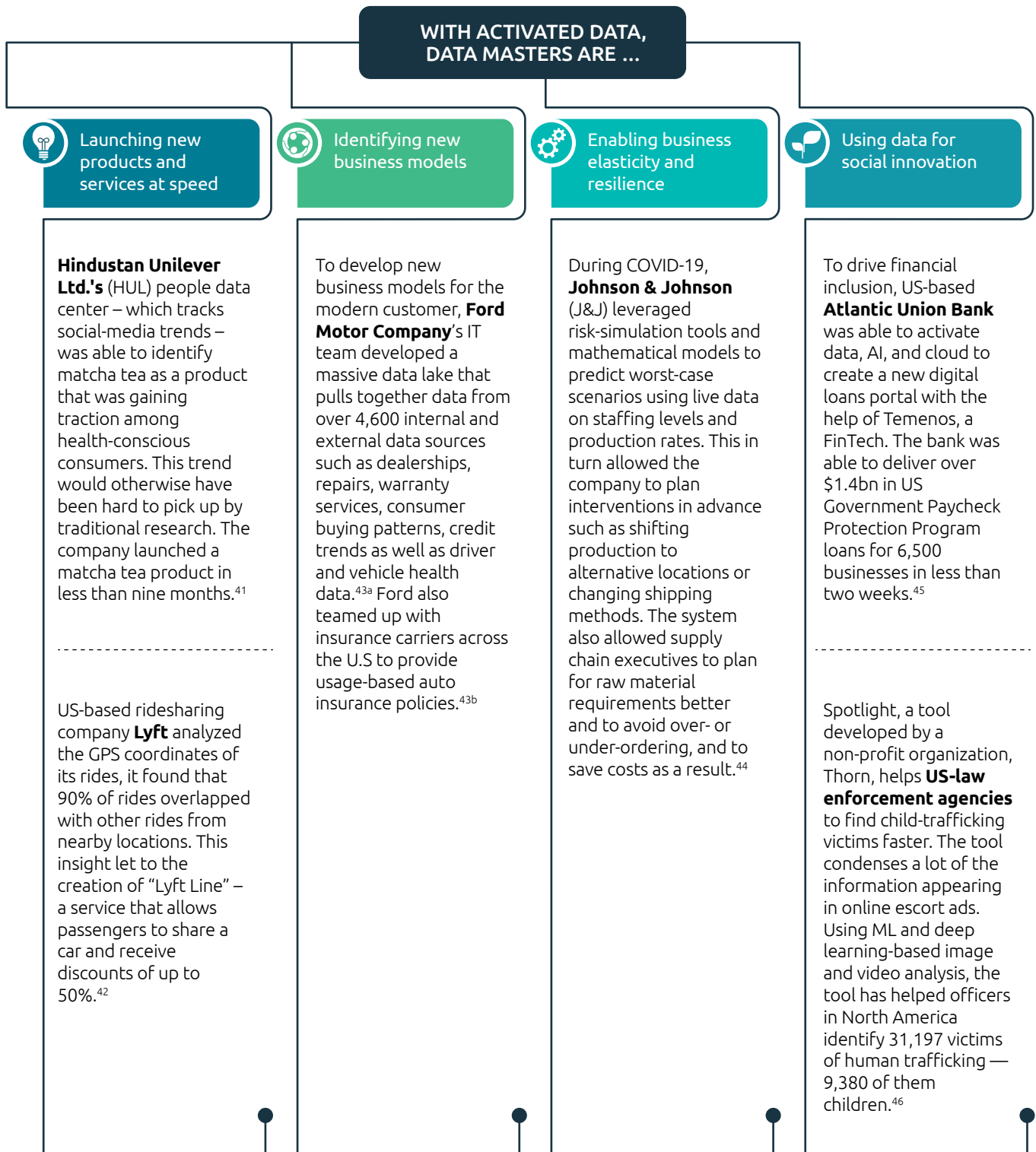


*Proximity with the business is fundamental because in the ocean of data, ‘boiling the ocean’ becomes a risk. So, focusing on areas most impactful for a user is key.”*

**RONAN CORRE**  
Enterprise data quality lead,  
Microsoft

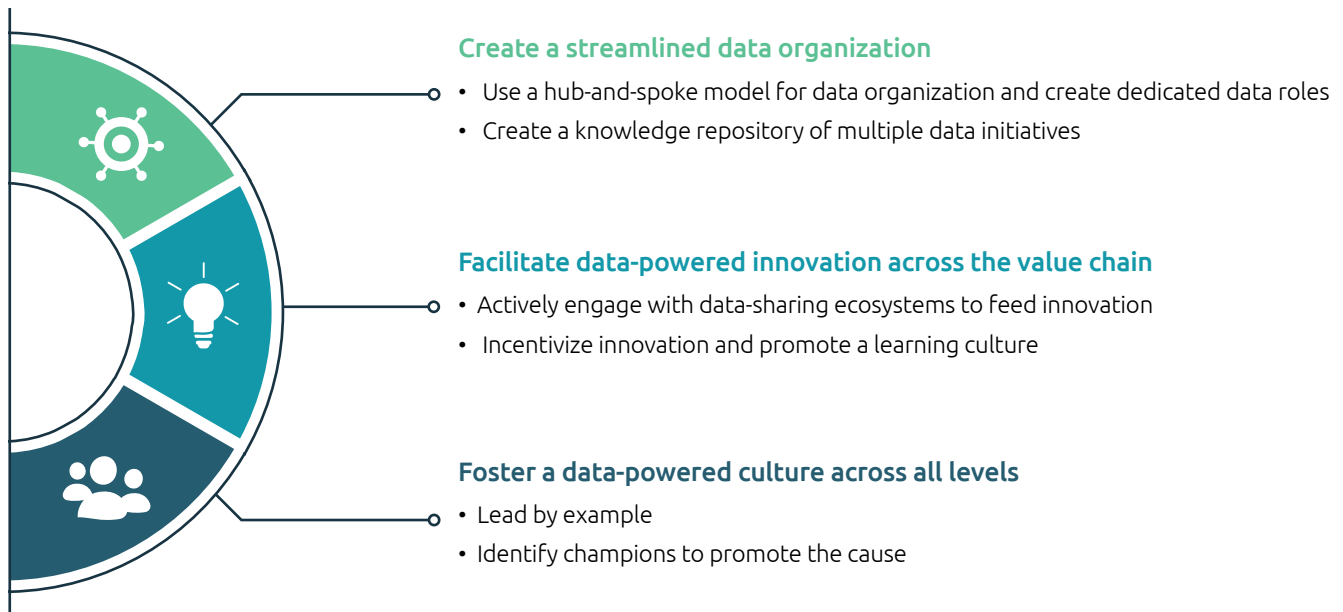


**FIGURE 5** Data masters are generating value with activated data and analytics



Source: Capgemini Research Institute Analysis; Company websites and news articles.

To activate data, analytics, and AI successfully, the following leading practices emerge among data masters:



Source: Capgemini Research Institute Analysis.

## 4.1 Data masters create a streamlined data organization

### WHY THE RIGHT DATA-OPERATING MODEL IS IMPORTANT?

A modern data platform will only bring value if the organization has processes in place to enable access and use of data at the right time to the right person. Accelerating speed of development will not suffice if data teams are not aligned on business objectives, or collaboration between IT and business teams or within data teams is missing.

The right operating model:

- Helps remove data silos between business units
- Ensures strong collaboration between data teams spread across the organization
- Creates strong linkages between data initiatives and overall corporate objectives.



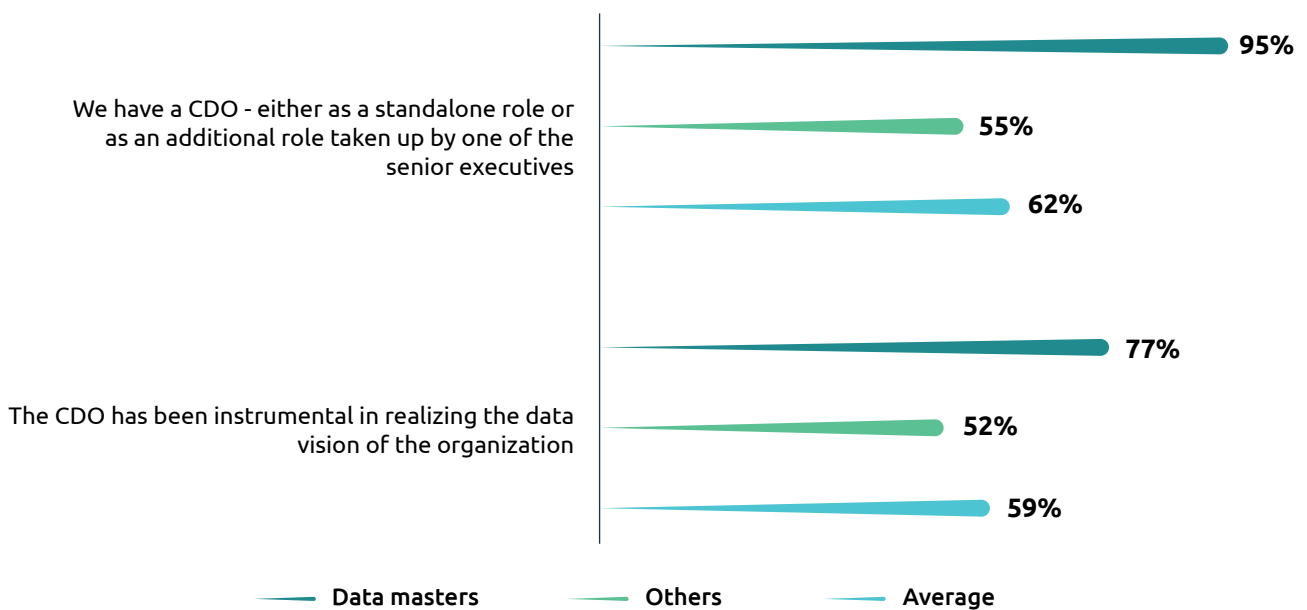
## How are data masters creating the right data-operating model?

### Data masters use a hub-and-spoke model for data organization and create dedicated data roles at leadership level

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

The operating model for data governance should evolve from a centralized model to a hub-and-spoke model where a chief data officer (CDO)-led central unit is responsible for policy making and governance, MDM and data quality, while the operationalization of initiatives is managed by the business units themselves. The data masters we surveyed earlier have one thing in common: 95% of them have a CDO, either as a standalone role (87%) or as an additional responsibility (8%), as Figure 6 shows. Essentially, members of these data organizations should be a mix of the business

**FIGURE 6** At data masters, the CDO plays a critical role in realizing the data vision



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

Data masters also tend to have a chief analytics officer (CAO) who is responsible for business intelligence, data science, and the activation of data. While a CDO is primarily responsible for managing and governing data, the CAO is responsible for using data. A CAO can report to the CDO or to the CEO/CIO even based on the scope of their responsibilities. A CDO or a CAO's role here is to provide guidance and direction to the organization while allowing the spokes to operate independently.

Effective data governance depends on several factors, including existing levels of centralization of decisions, organizational complexity, and maturity of data initiatives. Data masters primarily deploy a hub-and-spoke model to operationalize data initiatives.



# LEARNING FROM DATA MASTERS: THE HUB-AND-SPOKE APPROACH AT FACEBOOK

## Spoke:

Facebook has data analysts within all of its product teams who support the relevant department's use cases and issues related to specific products, as well as answering the team's data-related queries. Data analysts also build knowledge about the product they are assigned to and they can easily convert departmental requirements into required technical specifications.

## Hub:

Facebook also created a central team to coordinate the efforts of various data teams and share learning and best practices. The central team:

- Maintains the supporting infrastructure and provides access to the datasets required by data users
- Acts as the primary publisher of curated datasets, which are used by teams across the organization
- Ensures collaboration and data sharing among various data teams residing in product teams. The central team created a forum that allows all data professionals, data scientists, data engineers, and data analysts to discuss their data usage and objectives
- Enables Facebook data users to speak a common language and provides common definitions of data.
- At the corporate level, the CDO is given responsibility for nurturing a data-driven culture.<sup>47</sup>

## Data masters create a knowledge repository of multiple data initiatives

Translating data into action is a critical aspect of becoming data-powered. Data masters create a centralized platform for reviewing and sharing insights across the business.

Colgate-Palmolive built an insights platform to centralize the efforts of over 120 insights associates. The platform, powered by AI algorithms, checks new project briefs, prevents duplication of past research, and ensures that teams focus on new insights. "This allows us to shift our insights budget from tactical work to experimentation and foundational research that deepens our understanding of people and the world they live in," says Taylor Gordy, VP and GM Northern Europe at Colgate-Palmolive.<sup>48</sup>

Once a research project is completed, the platform automatically delivers the results to Colgate-Palmolive commercial managers and marketers all around the world, using targeted delivery channels. The platform also provides targeted insights based on business lines and regions, and AI recommendations based on the content they open, like, and share.<sup>49</sup>

## 4.2 Data masters facilitate data-powered innovation across the value chain

### WHY DATA-POWERED INNOVATION IS IMPORTANT?

The impact and value of data-powered innovation is only fully realized by achieving scale and widespread adoption, as opposed to ideation or experimentation.

For example, Colgate Palmolive has used AI and its database of over 80,000 oral-care formulas, as well as recent market trends, to reduce drastically the time to develop and market a new formulation. This has helped Colgate-Palmolive achieve major cost savings and gain competitive advantage in a crowded market.<sup>50</sup>



### How are data masters accelerating and scaling data-powered innovation?

Data masters realize the value of innovation and take active steps to enable and scale it. They define clear priorities and arrive at a prioritized use case roadmap. They also focus on building a supportive ecosystem, including startups and universities.

“We are building complex tools in order to mix key business variables to make life easier for innovation,” says Danone’s Ignacio Marinas. “So, from a consumer perspective, the tech team is trying to detect trends much more clearly and detect winners early on. Otherwise, it becomes too late for us to react to changes. And, for the different variables we have, we are trying to identify alarm signals and create a cockpit of innovation. This is where we mix competitor trends and consumer trends, and track innovation of what’s happening in the market already.”

**Data masters actively engage with data-sharing ecosystems to feed innovation:**

### WHY EXPLOITING EXTERNAL DATA IS IMPORTANT?

Data masters not only make best use of their internal data, but they also use external data to be more responsive, competitive, and innovative.

For instance, owing to pandemic-induced panic buying, US-based CVS Health could not extrapolate demand for some items from historical buying patterns. Taras Gorishnyy, vice president of enterprise retail analysis at CVS Health, says: “We needed to start learning very quickly from the changes that were happening in real time: what are the data elements that capture the new normal of the consumer behavior and how they have changed?”<sup>51</sup>

CVS started to find alternative datasets that could be incorporated as proxies into their demand models. These alternatives included real-estate data from Zillow; infection data from The Johns Hopkins Hospital and The New York Times; employment data from the Department of Labor; and news data from The GDELT Project.



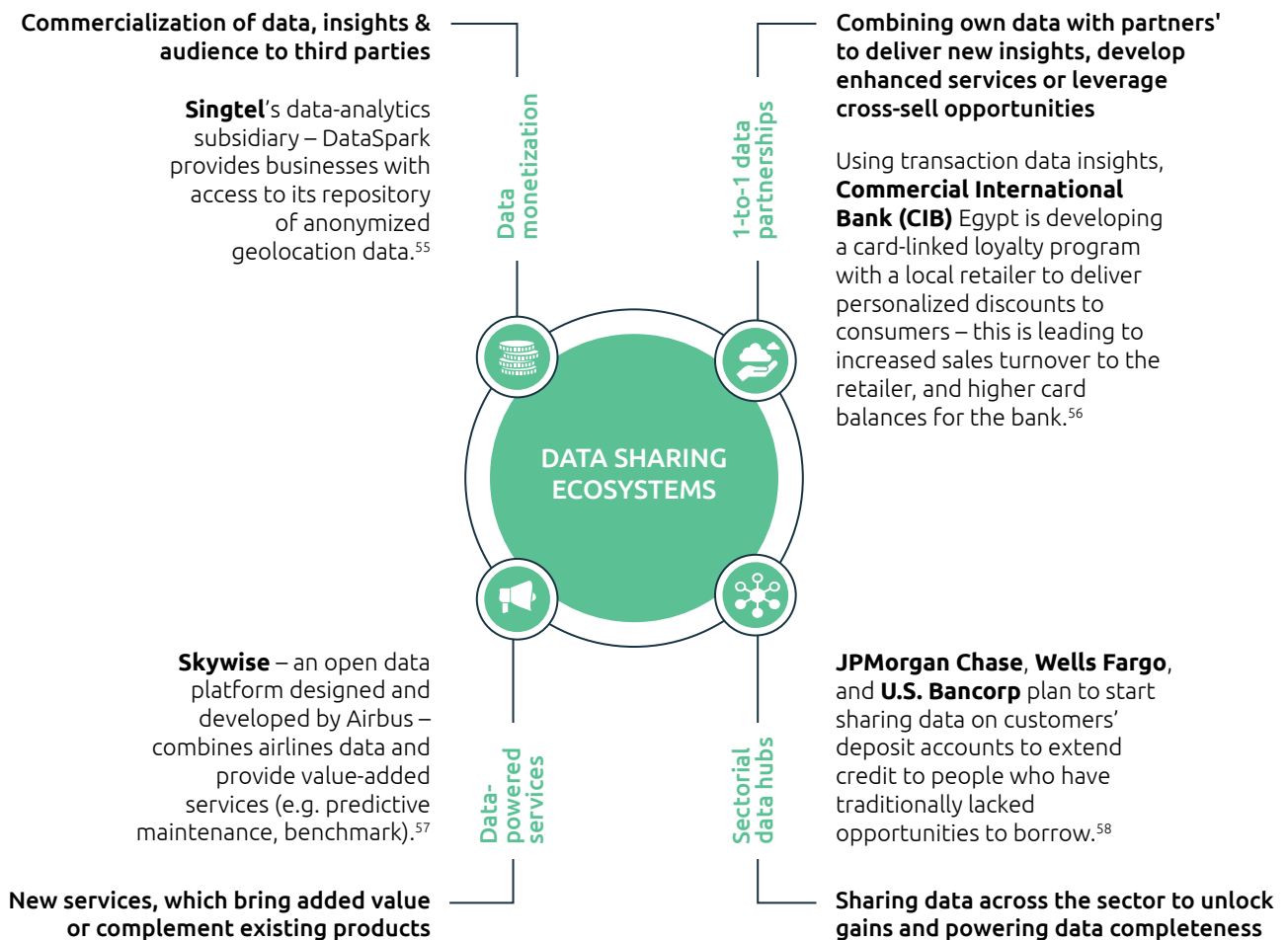
Surprisingly, the Zillow data turned out to be one of the most promising proxies for CVS. While the number of homes sold during “shelter-in-place” dropped, home prices largely held up. As time went on, CVS found a correlation between strong real-estate prices and customer buying behavior in specific CVS stores. Led by this data, CVS was able to predict demand within ±2%. Moreover, CVS used this data to create a set of “early warning” indicators of a likely surge or a dip in sales, enabling the company to take prescriptive action.<sup>52</sup>

Data masters are expanding their ability to capitalize on external data. But another option has emerged: data sharing and exchange. Companies such as BMW have been sharing their data to enable “data for good” applications. BMW recently partnered with Otonomo, an automotive data-services platform, to share its vehicle data to enable smart-city, real-time traffic information, and other solutions that improve the driving experience.<sup>53</sup> COVID-19 has also further demonstrated the need for businesses to work together to share data to help them innovate, achieve efficiencies, and build resilience.

Data masters have started to realize the value in data that is shared and collaborated on, all without risking their competitive advantage. Our previous research found that organizations participating in data ecosystems have the potential to gain financial benefits up to \$940 million or over 9% of annual revenue in the next five years.<sup>54</sup> Many models for data sharing have emerged:

**FIGURE 7**

**Data masters are looking at various data-sharing models to unlock new sources of revenue, introduce new services, and deliver new insights**



Source: Capgemini Research Institute Analysis; Company websites and news articles.

Data sharing does not necessarily come naturally to businesses, however. There are many barriers, from establishing mutual trust for proper handling of shared data, to understanding the sharing mechanisms, and ensuring compliance to regulations when sharing externally.

From our research, we were able to understand some best practices when accessing data ecosystems:

- **Prioritize data sharing use cases that are aligned with business goals:**

For instance, even before the COVID-19 outbreak, ten large pharmaceutical companies – including Johnson & Johnson, AstraZeneca and GSK – undertook collaborative efforts to train their ML algorithms on each other’s data to promote drug discovery. They used blockchain to share data without compromising confidential or commercial secrets. This was aligned with their target of accelerating the discovery of drugs and reducing the cost of the process.<sup>59</sup>

- **Explore various ways of data sharing:**

A recent survey reveals that 49% of organizations that commercialize data expose an API to that data and 46% sell an application that enables users to see data trends and insights.<sup>60</sup>

A former chief innovation officer at a credit insurance company says: *“We collect data on companies to assess the risk of default – and this data is a key asset to our business for taking decisions such as underwriting or selling insurance. But now we also extensively share this data to identify new value propositions and for monetization. We have created a product, called ‘TrafficLight,’ to provide an aggregated view on the financial health of a company.”*

- **Establish governance around sharing data:**

It is crucial to move away from a culture of ownership to one of data sharing. However, data masters also take into consideration the government and industry-related regulations regarding the security and privacy of data. Data masters also appoint “data-governance champions” for looking at these legislative mandates and setting up policies and procedures for effective data sharing.

- **Data masters incentivize innovation and promote a “learning” culture:**

- Innovative experiments do not always succeed, and a key component of an innovation culture is the ability to take risks, accept failure and learn from it. The idea of “failing fast” and “learning fast” is often discussed and encouraged in data masters.

- In data masters, innovation is not only encouraged, but also incentivized, and failure is embraced and celebrated. P&G gives out a “Heroic Failure” award each year to the individual or team who takes the greatest “intelligent” risk at the company. At the heart of the Heroic Failure award is an assurance to employees that they will not be punished for the mistakes they make, building a culture of “psychological safety” and trust.<sup>61</sup>



*We want to create this culture that quality data is a everybody’s matter. Everybody in the data supply chain needs to be accountable for that clean house because everybody benefits from it...”*

**DELPHINE CLEMENT**

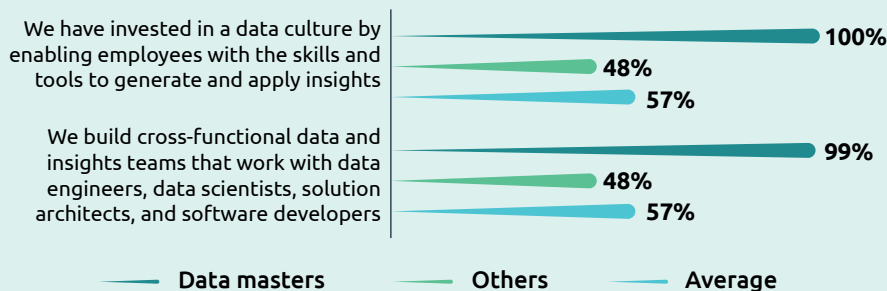
Director business program manager,  
Enterprise Data Management,  
Microsoft

## 4.3 Data masters foster a data-powered culture across all levels of management

### WHY DATA-POWERED CULTURE IS IMPORTANT?

In previous research we conducted, 62% of organizations said that establishing a digital culture was one of the top-two hurdles for digital transformation.<sup>62</sup> Further, our research on data-powered enterprises showed that almost all data masters invest in a data culture.

**FIGURE 8** Data masters invest in developing a data culture



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



### How do data masters build a data-powered culture?

#### Data masters lead by example

Data masters ensure data-driven behavior by all their team members. Roberto Giménez, digital business manager at Roche, says: “I have been working with people with clear leadership on data-driven decisions. Whenever they ask for resources or feedback, they also check whether this is based on data or based on emotions, thoughts, etc. When they have meetings, they ask that all the feedback that we bring to those meetings be supported by something, thereby creating this culture from the top down.”

#### Data masters identify champions among business teams who can promote their cause

A cultural change in the organization takes time and is often met with resistance from employees. In order to overcome

this resistance, data masters look to identify champions among the business teams who can promote their cause among their peers.

Camilla Schwartz-Björkqvist, global data & analytics transformation director at Danone, outlines how they identified key executives who were highly motivated by the objective to help lead the transition to a data-powered culture. “There’s a cultural transformation [in] a data and analytics journey,” she explains. “Firstly, we took a very conscious strategy of going with the front-runners. So, the executives that were excited to do this – who had either done it before or couldn’t wait to transform. We focused on them first – creating business cases together with them and making sure that we delivered ROI quickly so that we could use those success stories.”



# CONCLUSION

Organizations clearly grasp the value of data, but only a few are able to make data work for them. These fortunate few – our data-powered enterprises or the data masters – enjoy a significant business advantage. Organizations that want to emulate that advantage need to take a number of concrete steps:

1. **Data strategy** – The data masters have long since realized that business ownership of data transformation is a must-have. They work hard to establish a data strategy that is business-driven and owned. With that in place, they are then able to capitalize by the opportunities offered by combining both internal and external data.
2. **Data trust** – the data trust gap is marked, and data masters take concrete steps in reducing this trust deficit. They focus strong data management capabilities and data quality to build the base trust foundations in the veracity of data. They then focus on trusted AI systems and ethics as well as democratization of data.
3. **Modernization of data estate** – to ensure that the IT systems are able to support the increasing demand for data and analytics, data masters take a number of steps, including decommissioning legacy systems, migrating to multi-cloud environments, leveraging DataOps and developing customizable data-discovery tools.
4. **Activation** – this is the final but the critical step in becoming a data master. Data masters are successful in building a data organization that focuses on business priorities and is business-driven. A governance model that supports this data organization – coupled with a culture focusing on data-powered innovation – helps organizations get ahead of their peers.

With concerted and bold action in these four areas, organizations can join the data masters. And, by joining the ranks of the high performers, they will be better placed to drive growth, meet fast-changing customer needs, manage risk, and get ahead of shifts and volatility in their business environment.



*I have been working with people with clear leadership on data-driven decisions. Whenever they ask for resources or feedback, they also check whether this is based on data or based on emotions, thoughts, etc. When they have meetings, they ask that all the feedback that we bring to those meetings be supported by something, thereby creating this culture from the top down."*

**ROBERTO GIMÉNEZ**  
Digital business manager, Roche

# APPENDIX 1

## OUR DATA MASTERY FRAMEWORK

### 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

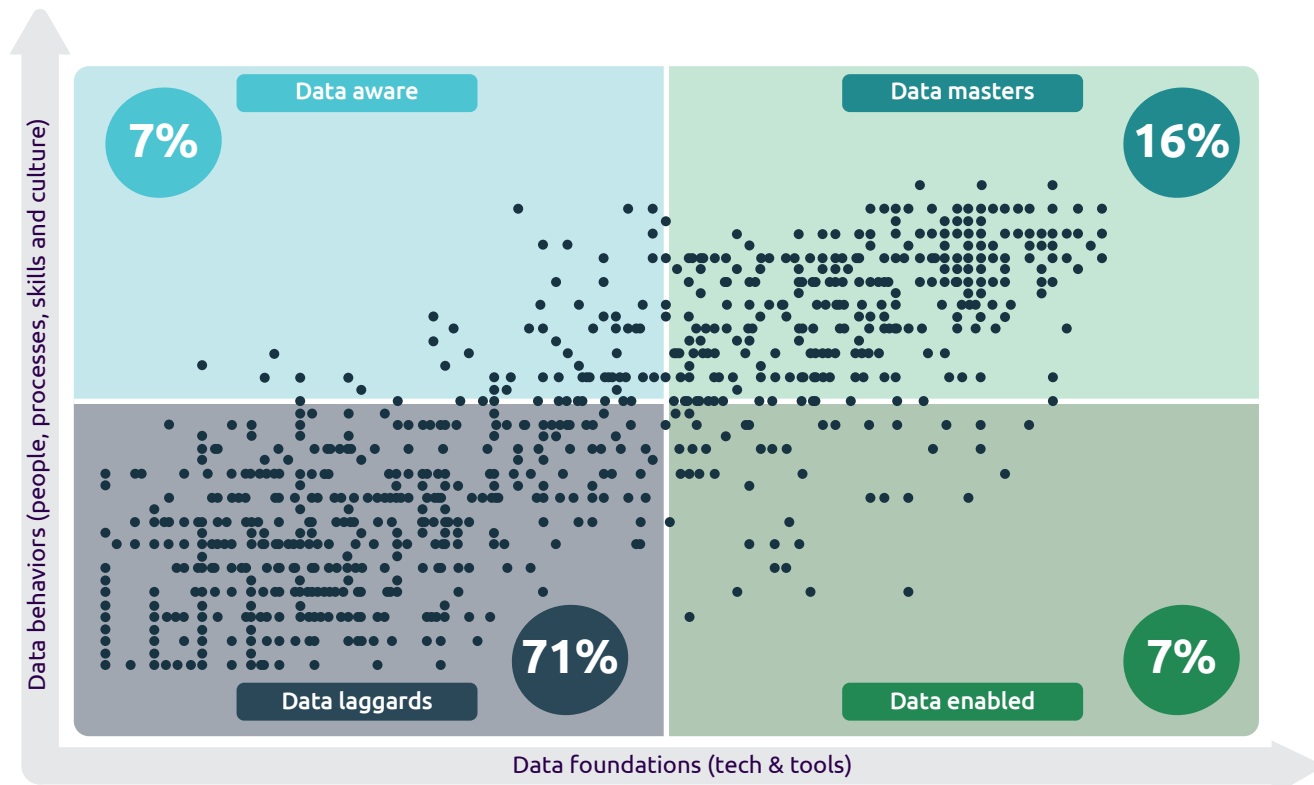


### Data behaviors

- 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



Only 1 in 6 organizations can be categorized as a data master



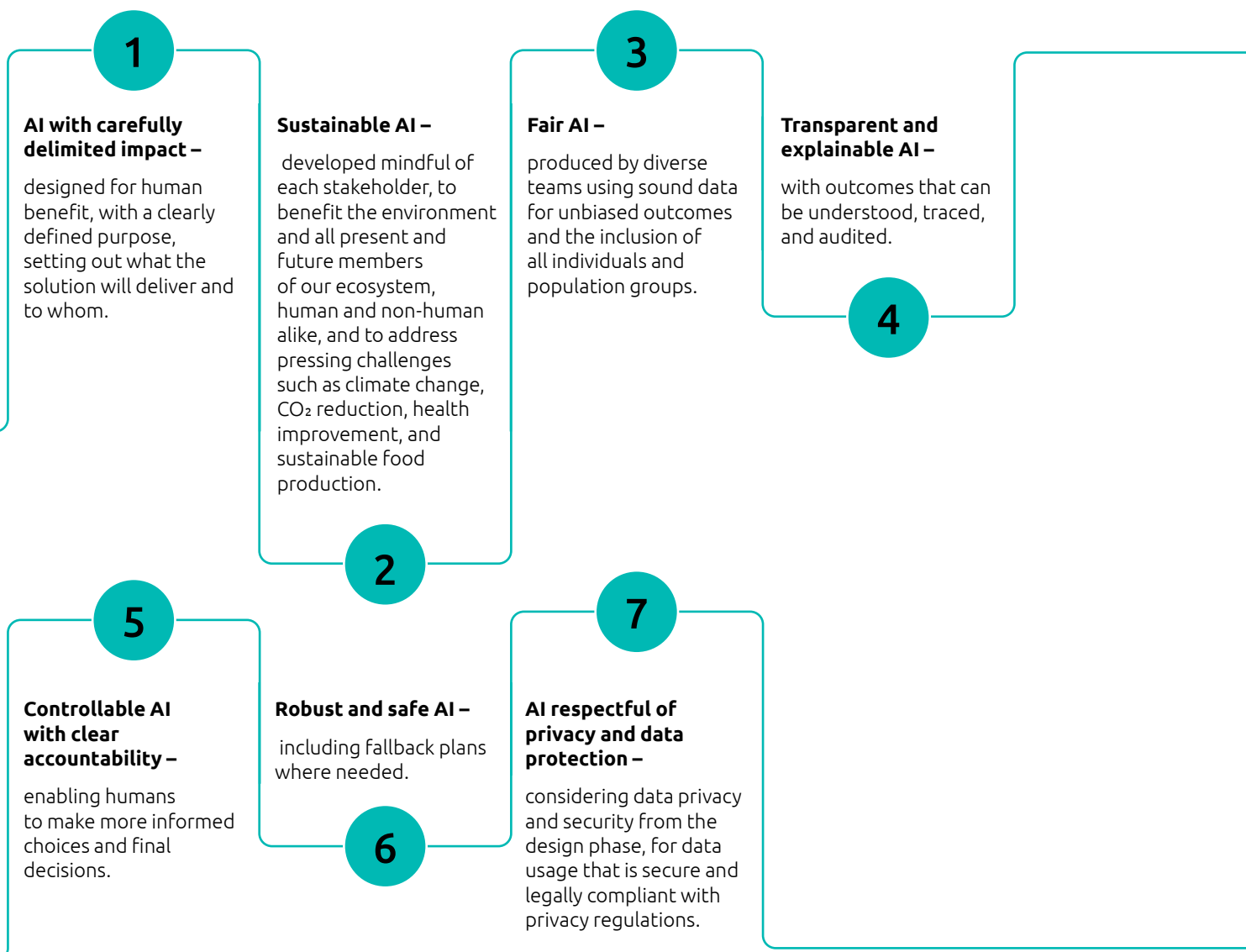
Source: Capgemini Research Institute analysis, Data-powered enterprises survey, August 2020; N=500 technology executives, N=504 business unit executives. Percentages are rounded off to the nearest integer and do not necessarily total 100.

# APPENDIX 2

## CAPGEMINI'S CODE OF ETHICS FOR AI

At Capgemini, our ethical culture drives our vision of AI. As a leader in digital transformation, we are committed to the adoption of AI in a way that delivers clear benefits within a trusted framework.

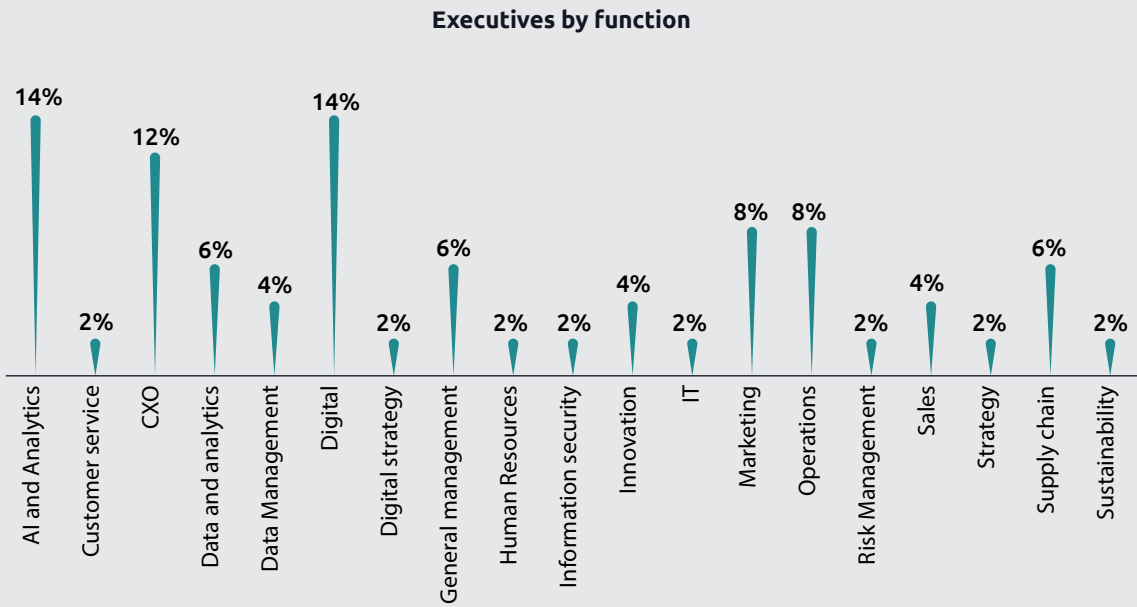
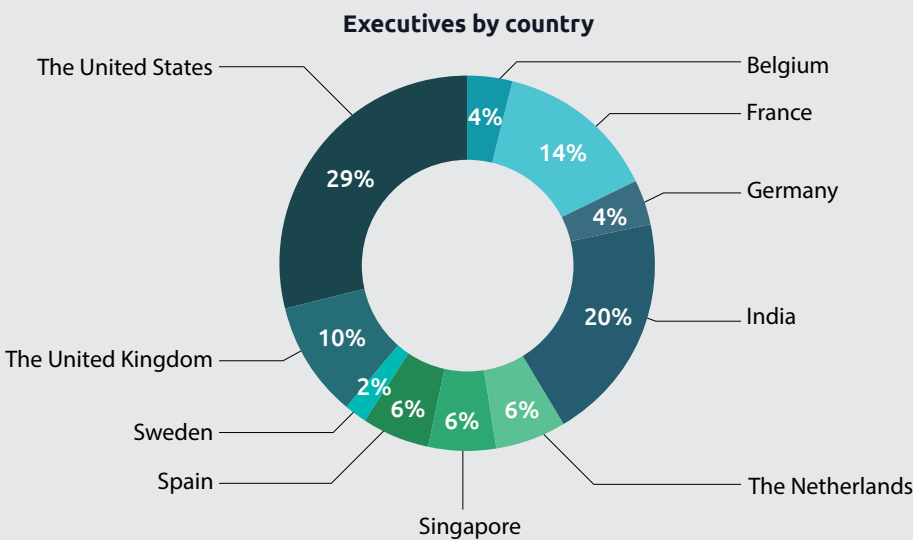
We believe that human ethical values should never be undermined by the uses made of AI by business. We want AI solutions to be human-centric, which we define as follows:



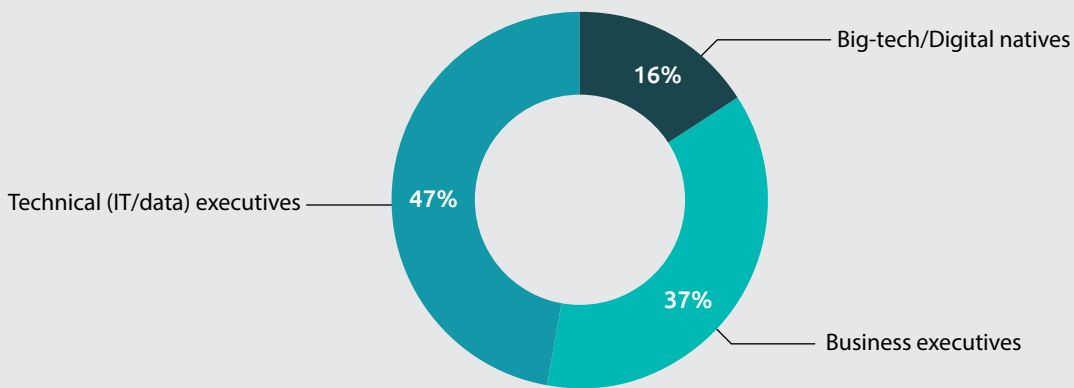
For more information, please visit: <https://www.capgemini.com/our-company/values-ethics/our-code-of-ethics-for-ai/>

# RESEARCH METHODOLOGY

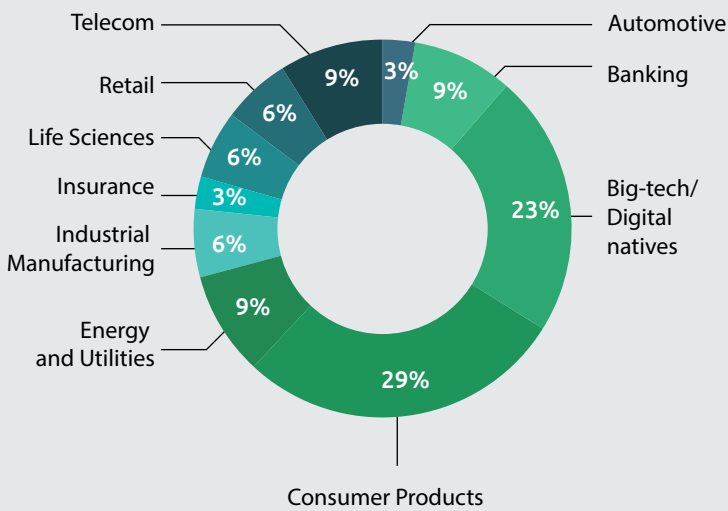
Between February to April 2021, we conducted in-depth interviews with 51 senior executives from 35 organizations with leading practices in data and analytics, across a range of sectors and countries. Two-thirds of these organizations reported revenues of more than \$20 billion for the last financial year. More than one-fifth of these organizations are either big-tech companies or digital natives.



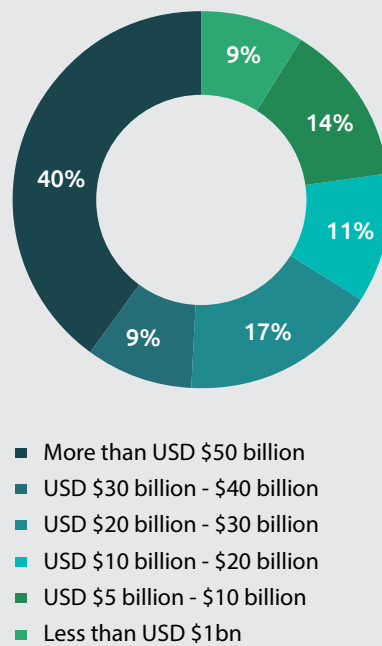
**Executives by company type**



**Organizations by industry**



**Organizations by revenue**



**Source:** Capgemini Research Institute, Learning from data masters in-depth interviews, N=51 executives from 35 leading organizations in data and analytics.

# REFERENCES

- 1 Capgemini Research Institute, "The data-powered enterprise: Why organizations must strengthen their data mastery," November 2020.
- 2 Ibid.
- 3 GDPR-info.eu, "Art. 5 GDPR: Principles relating to processing of personal data," accessed May 10, 2021.
- 4 Capgemini Research Institute, "DATA SHARING MASTERS: How smart organizations use data ecosystems to gain an unbeatable competitive edge," July 2021.
- 5 ESRI, "How GM maps and manages supply chain risk," November 2018.
- 6 ComputerWorld, "How Munich Re built a data lake fit for all its employees," April 2018.
- 7 Capgemini Research Institute, "DATA SHARING MASTERS: How smart organizations use data ecosystems to gain an unbeatable competitive edge," July 2021.
- 8 Capgemini Research Institute, "The data-powered enterprise: Why organizations must strengthen their data mastery," November 2020.
- 9 Gartner, "How to create a business case for data quality improvement," June 2018.
- 10 Pragmatic works, "The cost of bad data – infographic," May 2017.
- 11 Capgemini, "How data quality can hurt your data science programme... if you're not careful," February 2020.
- 12 Capgemini Research Institute, "The AI-powered enterprise: Unlocking the potential of AI at scale," July 2020.
- 13 Stanford University, "Artificial Intelligence index report," 2021.
- 14 Capgemini Research Institute, "AI and the Ethical Conundrum: How organizations can build ethically robust AI systems and gain trust," September 2020.
- 15 Capgemini Research Institute, "Why addressing ethical questions in AI will benefit organizations," September 2020.
- 16 Siemens blogs, "Achieving ethical AI for industrial applications," January 2021.
- 17 H&M group, "Meet Linda Leopold, Head of AI Policy," June 2019.
- 18 Reuters, "Amazon scraps secret AI recruiting tool that showed bias against women," October 2018.
- 19 Business Insider, "Google empowers 5,000 Cloud employees in ethical AI," July 2020.
- 20 Google Cloud, "Explainable AI," accessed May 28, 2021.
- 21 CFO, "The growing importance of data democratization," December 2019.
- 22 Forbes, "What is data democratization? A super simple explanation and the key pros and cons," July 2017.
- 23 Alation, "Munich Re launches enterprise-wide data-driven platform for analytics," October 2018.
- 24 Dataiku, "GE Aviation: from data silos to self-service."
- 25 Ibid.
- 26 CarMax analyst day transcript, May 6, 2021.
- 27 Shopify, "Deconstructing the monolith: designing software that maximizes developer productivity," February 2019.
- 28 DoorDash engineering, "Rebuilding our pricing framework for better auditability, observability, and price integrity," March 2021.
- 29 Diginomica, "Capital One closes its data centres and goes all in with AWS," January 2021.
- 30 Amazon Web Services website, "Nationwide selects AWS as its preferred cloud provider to deliver digital experiences to customers," December 2020.
- 31 Towards Data Science, "What is a Data Mesh — and How Not to Mesh it Up," July 2020.
- 32 Microsoft Azure, "Factory of the future with Microsoft Azure machine learning and AI," March 2020.
- 33 Intel Corp special call, April 06, 2021, via AlphaSense.
- 34 Talend, "Enabling customer-centric, data-driven marketing and sales."
- 35 Facebook Engineering, "Nemo: data discovery at Facebook."
- 36 Shopify Engineering, "How we're solving data discovery challenges at Shopify," July 2020.
- 37 TDWI, "How DataOps is transforming industries," April 2019.
- 38 Eckerson Group, "Best practices in DataOps: how to create robust, automated data pipelines," June 2019.
- 39 Streamset, "Humana - making healthcare easier for community members."
- 40 Yahoo Finance, "UPS slashes ground delivery transit times," November 2020.
- 41 Fortune India, "HUL: Eat. Sleep. Innovate. Repeat," December 2019.
- 42 AWS, "The data-driven enterprise."
- 43 a. AI Data and Analytics Network, "Ford's data-driven roadmap towards future mobility," February 2021.  
b. Ford Media Center, "Ford Supports InsureTech Revolution With Insurance Carriers, Ford Insure And Lincoln Motor Company Insure, Using Embedded Telematics To Help Customers Save On Auto Policies," December 2020.
- 44 Johnson & Johnson, "4 high-tech tools J&J is using to get products to you during the pandemic," September 2020.
- 45 World Economic Forum, "Cloud and SaaS technology can drive inclusive banking. Here are 3 reasons how," December 2020.
- 46 Amazon AWS, "Thorn uses AWS to help law enforcement identify child-trafficking victims faster," November 2017.
- 47 O'Reilly, "Building a data driven organization."
- 48 Marketlogic website, "Colgate Palmolive and Market Logic win 2021 BIG Innovation Award," January 2021
- 49 Cision, "Market Logic and Colgate-Palmolive win 2021 BIG innovation award," January 2021
- 50 Call transcript, "Colgate-Palmolive Co at Consumer Analyst Group of New York Conference," February 21, 2020.
- 51 Worldquant Predictive website, "Our CVS Health Predictions Project Featured in Datanami," March 2021
- 52 Datanami, "The data proxy that let CVS see around the COVID corner," March 2021.
- 53 Forbes, "BMW, Otonomo expanding smart city data-sharing plan," April 2020.
- 54 Capgemini Research Institute, "DATA SHARING MASTERS: How smart organizations use data ecosystems to gain an unbeatable competitive edge," July 2021.
- 55 DataSpark website, accessed May 2021.
- 56 Forrester, "Chief data officers: invest in your data sharing programs now," March 2021.
- 57 Airbus Skywise website, accessed May 2021.
- 58 Wall Street Journal, "JPMorgan, others plan to issue credit cards to people with no credit scores," May 2021.
- 59 Financial times, "Pharma groups combine to promote drug discovery with AI," June 2019.
- 60 Forrester Analytics Business Technographics Data And Analytics Survey, 2020, N=1,829 global data and analytics business decision-makers whose firms are commercializing data.
- 61 Forbes, "Are you embracing failure or encouraging failure?" June 2015.
- 62 Capgemini Research Institute, "The digital culture challenge: Closing the employee-leadership gap," 2017.

# ABOUT THE AUTHORS



## **Zhiwei Jiang**

CEO, Insights and Data Global Business Line  
[zhiwei.jiang@capgemini.com](mailto: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 data privacy, security, sustainability, and now more into AI ethics and social impact.



## **Anne-Laure Thieullent**

Managing Director, Artificial Intelligence & Analytics Group Offer Leader  
[annelaure.thieullent@capgemini.com](mailto:annelaure.thieullent@capgemini.com)

Anne-Laure leads Perform AI, Capgemini's Artificial intelligence & Analytics group offer. She works with clients to scale Artificial Intelligence, so it infuses everything they do, and they become AI-driven, data-centric and innovative. With 20 years of experience in big data, analytics and AI systems, from design to production roll-out, her passion is to bring companies what they need to transform themselves into intelligent enterprises. Anne-Laure is committed to guiding clients to increase activating their data, while cultivating the values of trust, privacy and fairness.



## **Ron Tolido**

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

Executive Vice President and Chief Technology Officer, Capgemini Insights & Data global business line. AI Futures domain lead for Capgemini's Technology, Innovation & Ventures (TIV) council. Lead author of the TechnoVision trend series.



## **Valérie PERHIRIN**

Managing Director, Data, Analytics and AI  
Capgemini Invent  
[valerie.perhirin@capgemini.com](mailto: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.



## **Marie-Caroline Baerd**

Executive Vice President, Data & AI – Financial Services, Capgemini Invent  
[marie-caroline.baerd@capgemini.com](mailto:marie-caroline.baerd@capgemini.com)

Marie-Caroline leads Capgemini Invent's Data & AI offer for the FS sector, developing thought leadership, new frameworks and solutions. She has a track record of more than 20 years in business management consulting, and is applying her experience in helping her clients leverage the full potential of Data & AI to innovate, enhance their revenues and efficiency, supporting their transformation journey from awareness and AI/Data strategy to implementation at scale within the organization.



## **Steve Jones**

EVP, Chief Data Architect  
[steve.g.jones@capgemini.com](mailto:steve.g.jones@capgemini.com)

Steve Jones is the Chief Data Architect for Capgemini's Group Portfolio. With over a decade in Big Data and Cloud, he's been at the front of helping our clients adopt new technologies to drive new business value and become more data driven. Steve was one of the first people to propose using mobile applications and analytics to help track and manage individuals carbon footprint and has continued to be at the front of Capgemini's CSR and Diversity initiatives.

**Jerry Kurtz**

Executive Vice President, Insights & Data  
[jerry.kurtz@capgemini.com](mailto:jerry.kurtz@capgemini.com)

Jerry has more than 30 years of IT and management consulting experience. He currently leads Capgemini's Insights and Data business in the Americas. During his career, Jerry has led a number of global enterprise transformation programs with a particular focus in Manufacturing, Hi-tech, Consumer Products, Retail, and Logistics industries. Some of his competencies include Big Data & Analytics, Artificial Intelligence, Internet of Things (IoT), Enterprise Transformation, SAP, Supply Chain Management, Shared Services, and Business Process Services. He has also lived and worked in Japan on two separate occasions for five years.

**Eric Reich**

Offer Leader and Global Head of AI & Data Engineering Insights & Data, Capgemini  
[eric.reich@capgemini.com](mailto:eric.reich@capgemini.com)

Eric is having 20+ years of experience in the Industry and is part of Capgemini Insights and Data Global practice where he leads globally AI & Data Engineering team and portfolio. With 16+ years of experience in the FS sector and 8 years+ in the data space, Eric had the opportunity to work with some of Fortune 500 customers in the Assessment, Design and Execution phases of their data driven transformation. He is also engaged in the development of new offers and solutions on the data, Analytics & AI space with the major cloud platform providers and ISV on the market.

**Jerome Buvat**

Global Head of Research and Head of Capgemini Research Institute  
[jerome.buvat@capgemini.com](mailto: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](mailto: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**

Assistant Director, Capgemini Research Institute  
[ramya.puttur@capgemini.com](mailto:ramya.puttur@capgemini.com)

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

**Gaurav Aggarwal**

Manager, Capgemini Research Institute  
[gaurav.aggarwal@capgemini.com](mailto:gaurav.aggarwal@capgemini.com)

Gaurav is a manager at the Capgemini Research Institute. He likes to assess how technology impacts businesses and understand how they respond to it. He is eager to learn about emerging business models, technologies, and trends across sectors.

The authors would like to thank Paul Hermelin, Dinand Tinholt, Maarten Deceuninck, Sandeep Bhatia, Gagan Mehta, Sarah McNeil, Subrahmanyam KVJ, and Mansi Gupta for their contribution to the research.

The authors would also like to thank Ashwin Yardi, Thomas Svahn, Nick Clarke, Yashowardhan Sowale, Chris Dickey, Monish Suri, Ingo Finck, Sergey Patsko, Christian Kaupa, Pierre-Adrien Hanaia, Isabell Schastok, Matthew Jones, Sandra Ficht, Seth Rachlin, Nilesh Vaidya, Rüdiger Eberlein, Giulia Carsaniga, Johanna Sundh, and Priyabrata Ghosal for their contribution to this research.

## About the Capgemini Research Institute

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

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



## For more information, please contact:

### Global

**Anne-Laure Thieullent**  
Vice President, Artificial Intelligence and Analytics Group Offer Leader  
[annelaure.thieullent@capgemini.com](mailto:annelaure.thieullent@capgemini.com)

**Valerie Perhirin**  
Managing Director in charge of Insight-Driven-Enterprise, Capgemini Invent  
[valerie.perhirin@capgemini.com](mailto:valerie.perhirin@capgemini.com)

**Steve Jones**  
Chief Data Architect  
[steve.g.jones@capgemini.com](mailto:steve.g.jones@capgemini.com)

**Eric Reich**  
AI & Data Engineering Offer Leader  
[eric.reich@capgemini.com](mailto:eric.reich@capgemini.com)

**Padmashree Shagrithaya**  
AI, Analytics & Data Science Offer Leader  
[padmashree.shagrithaya@capgemini.com](mailto:padmashree.shagrithaya@capgemini.com)

**Sébastien Guibert**  
Intelligent Process Automation Offer Leader  
[Sebastien.guibert@capgemini.com](mailto:Sebastien.guibert@capgemini.com)

### APAC

**Dr. Jing Yuan Zhao**  
AI Centre of Excellence Leader  
[jing-yuan.zhao@capgemini.com](mailto:jing-yuan.zhao@capgemini.com)

### France

**Marion Gardais**  
AI Centre of Excellence Leader  
[marion.gardais@capgemini.com](mailto:marion.gardais@capgemini.com)

### Germany

**Fabian Schladitz**  
AI Centre of Excellence Leader  
[fabian.schladitz@capgemini.com](mailto:fabian.schladitz@capgemini.com)

### Europe

**Robert Engels**  
AI Centre of Excellence Leader  
[robert.engels@capgemini.com](mailto:robert.engels@capgemini.com)

### India

**Chandrasekhar Balasubramanyam**  
AI Centre of Excellence Leader  
[chandrasekhar.balasubramanyam@capgemini.com](mailto:chandrasekhar.balasubramanyam@capgemini.com)

### North America

**Dan Simion**  
AI Centre of Excellence Leader  
[dan.simion@capgemini.com](mailto:dan.simion@capgemini.com)

### United Kingdom

**Joanne Peplow**  
AI Centre of Excellence Leader  
[joanne.peplow@capgemini.com](mailto:joanne.peplow@capgemini.com)

# EMPOWER YOUR BUSINESS LEVERAGING DATA & AI WITH CAPGEMINI

Data promises infinite possibilities. Now, more than ever, organizations need to activate its full potential. To become a **Data Master**, you must be at the helm, ready to harness data in critical dimensions, to outperform cohorts on financial performance and make crucial decisions from a qualitative base. A Data Master is a high-performer who is able to harness the full suite of data capabilities across its organization and ecosystem — turning insights into actions, and beyond.

Data Masters significantly outperform their peers, generating **+70% higher revenue** per employee and up to **22% higher profitability** and lead in top-line benefits increasing sales of traditional or new products and services.<sup>1</sup> More over front-runners who are participating in data ecosystems have the potential to gain financial benefits up to **\$940 million** or over 9% of annual revenue over the next five years.<sup>2</sup> There's an ocean of business opportunities to create new intelligent experiences, products and services leveraging data from your ecosystem.

## **Become a Data Master: Manage your data as a strategic asset**

Capgemini's Perform AI is a **portfolio of Data and Artificial Intelligence services** to help organizations take control of their data as a strategic asset and transform their business to be a leader in the data economy.

Our seamless approach, that connects business strategy and data operations, enable growth, performance, innovation with tangible benefits. We empower some of the largest data & AI transformations at scale in the market – activating Data & AI for CxOs. We improve organization's business and cultural change, data engineering and data science skills, and execute at scale with hybrid cloud Data & AI platforms and Intelligent Process Automation. Get ready to shift into data-mastery to sustain a continuous competitive advantage with trusted business outcomes.

- 1 Capgemini Research Institute, "The data-powered enterprise: Why organizations must strengthen their data mastery," November 2020.
- 2 Capgemini Research Institute, "DATA SHARING MASTERS: How smart organizations use data ecosystems to gain an unbeatable competitive edge," July 2021.

# PERFORM AI:

Activate data. Augment intelligence. Amplify outcomes.

Speed up your business outcomes with trusted Data & AI at scale

With **Perform AI**, we deliver high-performing, tangible business outcomes, 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 product 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 assistance 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.

\*Capgemini wins 2020 Artificial Intelligence Breakthrough award for best virtual agent solution



# CAPGEMINI PERFORM AI

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

<b>CP &amp; RETAIL</b> Augmented operations and immersive customer engagement	<b>FINANCIAL SERVICES &amp; INSURANCE</b> Right Product, Right Price	<b>MANUFACTURING, AUTO &amp; LIFE SCIENCES</b> Augmented operations for operational efficiency & risk reduction
--	---	--

<b>TELCO MEDIA TECH</b> Deep Customer Engagement and Operational Excellence	<b>ENERGY UTILITIES &amp; CHEMICALS</b> Sustainability powered by Data & AI and Operational Excellence	<b>PUBLIC SECTOR</b> Infusing AI to better serve citizens and governments
--	---	--

## DATA ACTIVATION

## DATA ADVANTAGE

<b>CUSTOMER FIRST</b>	<b>INTELLIGENT INDUSTRY</b>	<b>ENTERPRISE MANAGEMENT</b>	<b>DATA SHARING ECOSYSTEMS</b>
-----------------------	-----------------------------	------------------------------	--------------------------------

## DATA ENABLERS

### Strategy & MVPs

**AI ACTIVATE**

Transformation Target and foundations for execution of AI @Scale



**AI REIMAGINE**

Invent and Implement "the next" Data & AI powered business

### Build, Deploy, Manage, Operate

**AI, ANALYTICS DATA SCIENCE** 890 - Capgemini

Scalable, Trusted AI and Augmented BI & Data Viz solutions

**INTELLIGENT PROCESS AUTOMATION**

Data & AI powered Frictionless Enterprise

**AI & DATA ENGINEERING** IDEA - Capgemini

Hybrid Cloud Data & AI platforms, Data Estate modernization, Intelligent Master Data Management & Data Trust



AI4Good



Trusted AI



Sustainable AI

Discover more about our research:



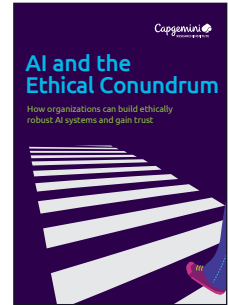
The data-powered enterprise:  
Why organizations must strengthen their data mastery



Data sharing masters:  
How smart organizations use data ecosystems to gain an unbeatable competitive edge



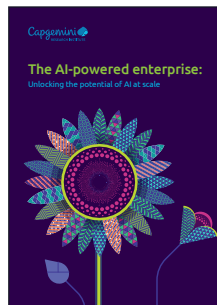
The Age of Insight:  
How Consumer Products and Retail organizations can accelerate value capture from data



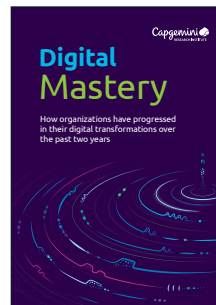
AI and the Ethical Conundrum:  
How organizations can build ethically robust AI systems and gain trust



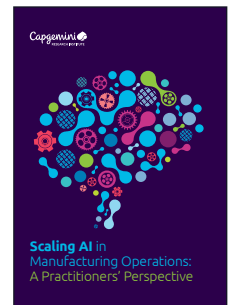
The Art of customer centric Artificial Intelligence  
Raising the bar on AI in customer experience



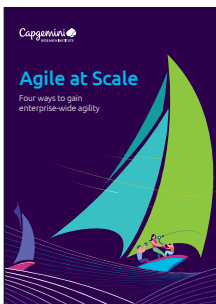
The AI-powered enterprise  
Unlocking the potential of AI at scale



Digital Mastery:  
How organizations have progressed in their digital transformations over the past two years



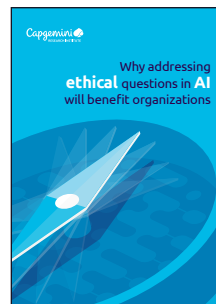
Scaling AI in Manufacturing Operations:  
A Practitioners' Perspective



Agile at Scale  
Four recommendations to gain enterprise-wide agility



Scaling Innovation:  
What's the Big Idea? Why most innovations fail to scale and what to do about it



Why addressing ethical questions in AI will benefit organizations



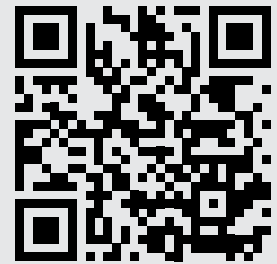
The Digital Culture Challenge:  
Closing the Employee-Leadership Gap



# 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/>

A screenshot of a web form titled "Capgemini Research Institute". It includes a close button in the top right corner. Below the title, it states "Fields marked with an \* are required". There are three input fields for "First Name", "Last Name", and "Email". Below these fields is a checkbox with the text "By submitting this form, I understand that my data will be processed for the purposes indicated above and described in the Terms of use." and a "Submit" button at the bottom.



## About Capgemini

Capgemini is a global leader in partnering with companies to transform and manage their business by harnessing the power of technology. The Group is guided everyday by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of 270,000 team members in nearly 50 countries. With its strong 50 year heritage and deep industry expertise, Capgemini is trusted by its clients to address the entire breadth of their business needs, from strategy and design to operations, fuelled by the fast evolving and innovative world of cloud, data, AI, connectivity, software, digital engineering and platforms. The Group reported in 2020 global revenues of €16 billion.

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