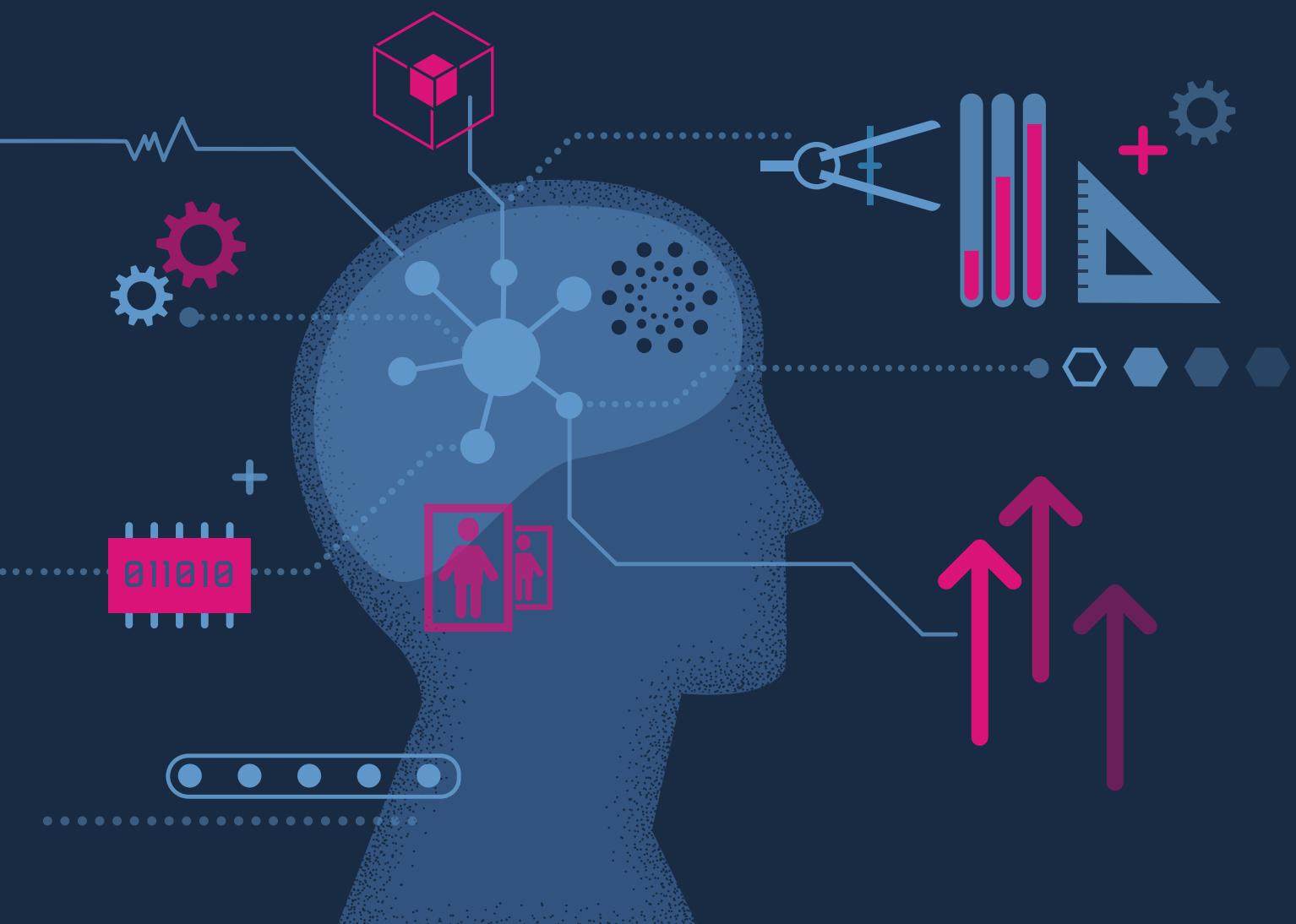


CIO vision 2025: Bridging the gap between BI and AI



Preface

"CIO vision 2025: Bridging the gap between business intelligence (BI) and AI" is an MIT Technology Review Insights report sponsored by Databricks. To produce this report, in May and June 2022, MIT Technology Review Insights conducted a global survey of 600 chief information officers, chief technology officers, chief data and analytics officers, and other senior data and technology executives. We also interviewed 10 C-level executives from Fortune 500 companies and successful start-ups. The survey respondents are evenly distributed among North America, Europe, and Asia-Pacific. There are 14 sectors represented in the sample, and all respondents work in organizations earning \$500 million or more in annual revenue. The research also included a series of interviews with executives who are directly involved in their organizations' AI and machine learning initiatives. Denis McCauley was the author of the report, Francesca Fanshawe was the editor, and Nicola Crepaldi and Natasha Conteh were the producers. The research is editorially independent, and the views expressed are those of MIT Technology Review Insights.

We would like to thank the following executives for providing their time and insights:

Sherry Ahom, Chief Digital Officer, Cummins

Vittorio Cretella, Chief Information Officer, Procter & Gamble

David Hogarth, Chief Information Officer, Virgin Australia

Marc Kermisch, Chief Information Officer, CNH Industrial

Swamy Kocherlakota, Chief Information Officer, S&P Global

Mike Maresca, Global Chief Technology Officer, Walgreens Boots Alliance

Masashi Namatami, Group Chief Digital Officer, Managing Executive Officer, Tokio Marine

Jeremy Pee, Chief Digital and Data Officer, Marks & Spencer

Prasad Ramakrishnan, Chief Information Officer, Freshworks

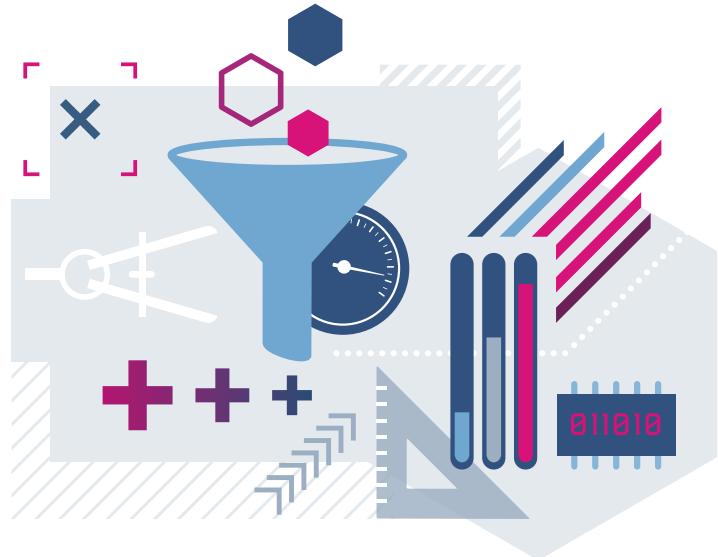
Rowena Yeo, Chief Technology Officer & Global Vice President, Technology Services, Johnson & Johnson



CONTENTS

01	Executive summary	3
	About the survey	4
02	Room to grow with AI	5
	Lofty ambitions	5
	Tokio Marine: Striving to become AI-driven.....	7
	Databricks perspective.....	8
03	A shift to financial value realization.....	9
	AI use case development to 2025:	
	Selected company examples.....	10
04	Meeting the challenges of scale.....	11
	Procter & Gamble (P&G): Automating to scale	13
05	The data priorities.....	14
	Priorities in focus	15
	Multi-cloud and open	17
	CNH Industrial: AI, open data, and the sustainable tractor	18
	An industry lens on data and AI.....	19
06	Conclusion.....	20

01 Executive summary



It's been several years since organizations began adopting artificial intelligence (AI) to improve their business; few have come close to mastering its existing capabilities. A small number of organizations in our research aim to become AI-driven – a status we define as AI and machine learning underpinning almost everything the enterprise does – by 2025. However, this elite group – who we term "AI leaders"¹ – as well as the many others looking simply to embed AI more firmly in the enterprise foundations face formidable challenges to achieving their objectives.

Addressing shortcomings in companies' data management and infrastructure, as well as internal structural and process rigidities and talent deficits, loom large among those challenges. Seventy-two percent of the technology executives we surveyed for this study say that, should their companies fail to achieve their AI goals, data issues are more likely than not to be the reason. Improving processing speeds, governance, and quality of data, as well as its sufficiency for models, are the main data imperatives to ensure AI can be scaled, say the survey respondents.

This report sheds light on these and other data constraints that organizations must address to unleash the potential AI holds for their businesses.² It also identifies the investments and other measures companies plan to take to align their data capabilities more closely with their AI ambitions. The study's findings are based on a global survey of 600 chief information officers, chief technology officers, and other senior technology leaders. We also drew insights from in-depth discussions with 10 such executives.

Following are the study's key findings:

- **Companies view wider AI adoption as mission-critical for their future.** From mostly limited AI use across the enterprise today, the surveyed executives plan a major expansion of use cases in all core functions in the next three years. Well over half expect AI use to be widespread or critical in their IT, finance, product development, marketing, sales, and other functions by 2025. While most will pursue a wide variety of use cases, many also aim to boost AI's impact on the top line, increasing the returns from revenue-generating uses.
- **Scaling AI successfully is priority one for the data strategy.** The surveyed companies' data and AI strategies are closely interlinked. Over three-quarters (78%) of the executives we surveyed – and almost all (96%) of the leader group – say that scaling AI and machine learning use cases to create business value is their top priority for enterprise data strategy over the next three years.
- **Major spending growth is planned to bolster AI's data foundations.** The surveyed CIOs – especially those in the leader group – plan sizeable increases in investment between now and 2025 to strengthen different parts of their data and AI foundations. Leader spending on data security over the next three years will rise by 101%, on data governance by 85%, on new data and AI platforms by 69%, and on existing platforms by 63%. (The analogous figures among the sample as a whole are 59%, 52%, 40%, and 42%, respectively.)
- **Investment growth intentions are strongest in the financial services industry.** Among the 14 industries in

the survey, AI leaders are most numerous among retail/consumer goods and automotive/manufacturing companies. Expected investment growth in these sectors in the above areas of data management and infrastructure is higher than in others with one exception: planned increases by financial service providers will substantially exceed those in all other sectors.

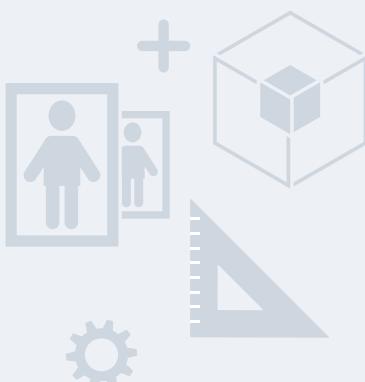
- **Multi-cloud and open standards are integral to AI progress.** Most of the survey respondents (72%) – and almost all leaders (92%) – appreciate the flexibility that a multi-cloud approach provides for AI development. CIOs interviewed for the study also emphasize the role of open architecture standards in supporting multi-cloud, and the importance of both in progressing their AI development.

About the survey

The survey that forms the basis of this report was conducted by MIT Technology Review Insights in May and June 2022. Following are the key demographic details of the 600 executives who took part in it.

The respondents hold senior technology roles in their organizations. The majority (84%) are C-level executives: chief information officers, chief technology officers, chief data/analytics officers, and chief AI officers (CIOs are 72% of the total sample). The balance consists of senior vice-presidents or vice-presidents of AI, of data platforms, or of engineering, and heads of AI and machine learning.

These executives work in predominantly large organizations. While 10% of the latter earn annual revenue of between \$500 million and \$1 billion, 45% earn between \$1 billion and \$5 billion and the other 45% – \$5 billion or over. Just over three-quarters (76%) employ over 5,000 people.



In terms of geography, North America accounts for 35% of the respondents, with the rest divided equally between the other two regions.

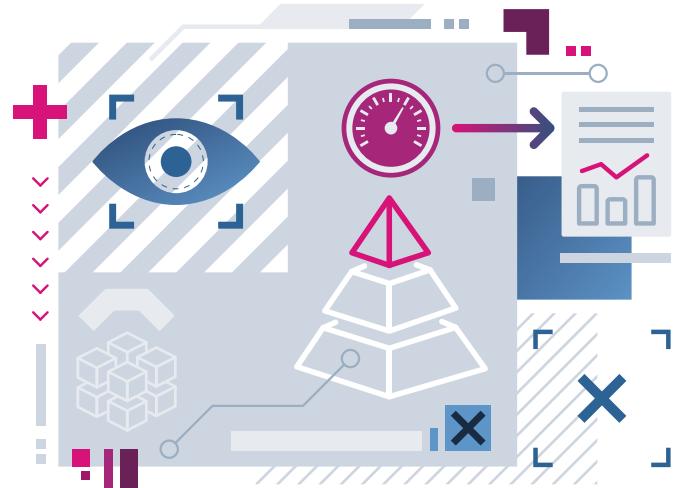
Eighteen countries are represented:

North America	Europe	Asia-Pacific
Canada	Belgium	Australia
United States	Denmark	India
	France	Japan
	Finland	Singapore
	Germany	South Korea
	Iceland	
	Luxembourg	
	Netherlands	
	Norway	
	Sweden	
	United Kingdom	

A total of 14 industries are represented in the survey:

Automotive/manufacturing	Media/entertainment
Education	Oil and gas
Financial services	Power and utilities
Government/public sector	Professional services
Life science and healthcare	Real estate and construction
Logistics/transport	Retail/consumer goods
	Technology
	Telecommunications

02 Room to grow with AI



Nearly a decade after they emerged from science labs, AI and machine learning are firmly embedded in enterprise technology environments and are starting to generate value for many businesses. It is increasingly difficult to find organizations that have not at least explored AI use in their business in some way. In our survey of 600 CIOs and other technology leaders, the share of those saying their companies are not using AI today is 6% or less in any of seven core enterprise functions (Figure 1, next page).

Although the hype surrounding AI and machine learning has largely subsided and use case development is widespread, these technology fields – and especially their commercial application – are still early in their maturity.³ The majority of survey respondents claim no more than limited adoption of AI uses today in all but two core enterprise functions, the exceptions being AI use by IT

The majority of survey respondents claim no more than limited adoption of AI uses today in all but two core enterprise functions, the exceptions being AI use by IT and by finance.

and by finance. Organizations have only scratched the surface of what such capabilities can deliver.

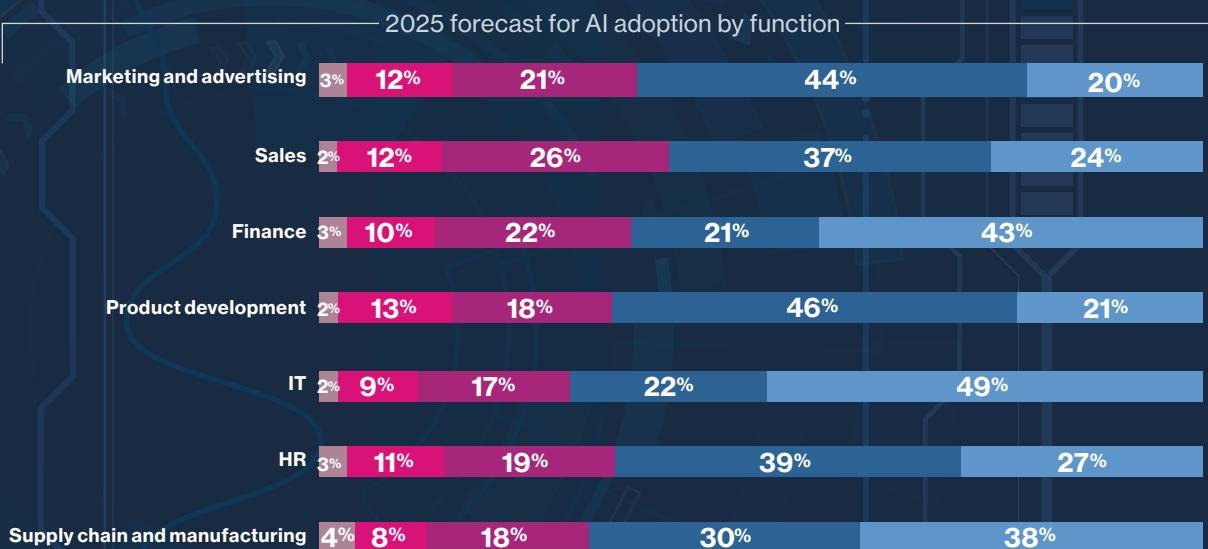
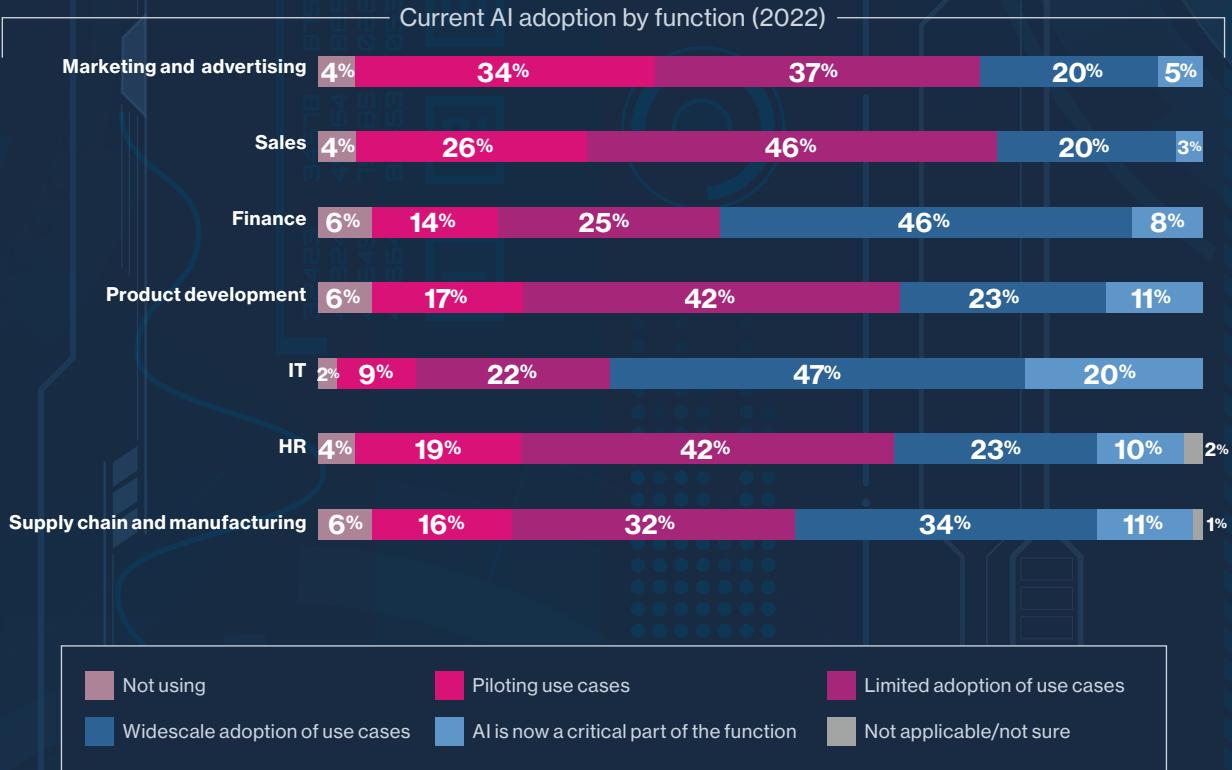
Less than 1% of the respondent companies can be considered AI-driven today, if that status is defined as AI being intrinsic to everything the organization does across most of its core functions. A select group of 14% – termed by us as “AI leaders” – aim to achieve that status by 2025, however, planning for AI to become “a critical part” of at least five core functions by then.

Lofty ambitions

The “AI leaders” are not alone in having ambitious plans for further AI adoption. In the survey sample as a whole, the share of respondents expecting AI use to be widespread or critical in 2025 range from 61% in sales, to 67% in product development, to 71% in IT. Even for today’s laggards, the percentage not using AI will drop by nearly half across all business functions. All this presages a significant expansion in the number of use cases organizations are developing.

Among the companies participating in our study, a few lay credible claim to being led by AI across their operations today. One is Freshworks, a US software-as-a-service (SaaS) provider of customer service and employee support solutions, whose business model with customers is largely delivered through AI using conversational messaging. “We started integrating it very early in the life of our business,” says Prasad Ramakrishnan, the company’s chief information officer. “When the company was founded [in 2011], we knew that AI would be a game-changer.”

Figure 1: The extent of AI use in core functions of the business, today and expected in 2025
 (% of respondents)



Another is S&P Global, a financial information and analytics company. Since acquiring an AI solutions provider in 2018, “AI, machine learning, and natural language processing have become embedded in everything that we do,” according to Swamy Kocherlakota, S&P Global’s chief information officer.

For most organizations in our study, however, becoming AI-led is a work in progress. “We’ve been aggressive in using AI to transform the digital experiences for customers in our omnichannel network,” says Jeremy Pee, chief digital and data officer at retailer Marks &

Spencer. “But we need to leverage AI to make ourselves better in every way. So, we are starting to use it in the core of how we run the business, how we make decisions, and how we put intelligence and science into that,” says Pee.

According to Masashi Namatame, group chief digital officer and managing executive officer of Tokio Marine, a Japanese insurance provider, becoming AI-driven means “applying AI as broadly, as aggressively, and as enthusiastically as possible. No part of our business should be untouched by it.”

Tokio Marine: Striving to become AI-driven

AI has already become deeply integrated into the insurance business. Insurers of all types now routinely use AI models to drive underwriting, streamline claims processing and accelerate claims adjudication, protect against insurance fraud, and improve risk forecasting, for example.

It is proving a source of disruption in some markets, as insuretech start-ups use their native AI capabilities to challenge established providers. And many of the latter are responding, using the vast troves of historical data at their disposal to develop impactful use cases of their own.

The experience of Tokio Marine – Japan’s oldest insurance company, founded in 1879 – offers a glimpse into the benefits AI deployment offers to established insurers as well as the challenges they face in mastering its use. “We are striving to become an AI-driven company,” says Masashi Namatame. “We are still in the process of learning from AI and trying to find more and better ways of applying it in our business.”

The use of AI is well advanced in Tokio Marine’s claims operations, and particularly in its auto insurance business, says Namatame. To assess collision damages, the company uses an AI-based computer vision solution to analyze photos from accident scenes. Comparing these with what he describes as “thousands or even millions” of photos of past analogous incidents, the model produces liability assessments of the parties involved and projects

anticipated repair costs. AI has also provided the company with tangible benefits in online sales – especially in personalized product recommendations – and in contract writing, according to Namatame.

Use cases currently in development include the analysis of data from in-car drive recorders, which monitor driver actions and behaviors. Such models, according to Namatame, will help further refine policy underwriting as they project the future risk of collisions posed by individual drivers. Improving fraud detection with AI is another priority for the company, he says.

In property insurance, photo recognition figures in an emerging AI use case that Namatame has high hopes for – mitigating climate change risk. He explains: “Existing claims assessment procedures conducted by humans are extremely time-consuming and dangerous when it comes to typhoons, flooding, and other natural disasters. We are now looking to feed drone and satellite data into our models to assess claims from such events.”

Namatame acknowledges the constraints that insurers like Tokio Marine face in scaling AI. Among these is the challenge of rendering historical data in the company’s legacy systems “fully AI-friendly”, and that of properly integrating external data into its AI models. Just as critical, Namatame adds, is overcoming the cultural challenges involved: “In order to become AI-driven,” he says, “we need to change the mentality of our entire business.”

Databricks perspective

Unify and scale your data warehousing and AI use cases on a single platform

Companies worldwide are eager to tap the potential of AI to increase innovation and efficiency. Based on this survey, 94% of companies are adopting AI in some capacity today. However, only 14% said they aim to be AI-driven by 2025. CIOs cite future-proofing data and AI foundations, investing in the “right” use cases that maximize ROI, and scaling effectively by leveraging multi-cloud, open standards, and open data as the three key strategies. What is holding back so many leaders from executing these strategies?

The challenge starts with the data architecture. Organizations need to build four different stacks to handle all of their data workloads: business analytics, data engineering, streaming, and machine learning (ML). All four of these stacks require very different technologies and, unfortunately, they sometimes don’t work well together. The result is multiple copies of data, no consistent security/governance model, closed systems, and less productive data teams. Meanwhile, ML remains an elusive goal. With the emergence of the lakehouse architecture, organizations are no longer bound by the confines and complexity of legacy architectures. The lakehouse architecture provides flexible, high-performance analytics, data science, and ML by combining the performance, reliability, and governance of data warehouses with the scalability, low cost, and workload flexibility of the data lake.

The Databricks Lakehouse Platform unifies and scales data, analytics, and AI capabilities in the following ways:

- **Multi-cloud:** Databricks is the only unified data platform across all three major public clouds (AWS, Azure, Google Cloud), meaning one tool for data engineering, data science, ML, and analytics. We also offer Databricks technology with Delta Lake in China by partnering with Alibaba.
- **Open:** We make it open to avoid lock-in by utilizing open standards and open data access and tapping into innovation from the open source community.

Companies worldwide are eager to tap the potential of AI to increase innovation and efficiency.

- **High performance, low cost:** Databricks Delta Lake dynamically changes the size of data partitions for the best combination of cost and performance. Databricks SQL allows customers to operate a multi-cloud lakehouse architecture that provides up to 12 times better price/performance than traditional cloud data warehouses.

- **Scalable and collaborative:** Our Data Science and Machine Learning platform allows developers and data scientists to explore their data, build, and productionize models, and share their analyses at scale. With an automated full ML lifecycle, you can shorten time from experimentation with ML models to robust production deployments.

As of July 2022, more than 7,000 customers globally and more than 50% of Fortune 500 companies use Databricks. Millions of machines are launched daily, hundreds of thousands of data scientists log on each month, and multiple exabytes of data are processed each day with Databricks Lakehouse. Our robust ecosystem encompasses over 500 consulting partners, over 100 ISV partners, and over 400,000 users from 150,000 companies for the free Community Edition of Databricks. Choosing the right technology platform and partner is the key that opens the doors to scaling data and AI. Databricks has proved leadership across the data and AI lifecycle. In fact, Databricks is the only cloud-native vendor named a Leader in both 2021 Gartner Magic Quadrants for Cloud Database Management Systems and Data Science and Machine Learning Platforms.

03

A shift to financial value realization

An expansion of use cases in production is one indicator of AI's growing impact, but ultimately the more important determinant are the types of value – and their magnitude – that it is delivering to the organization. "We've got several hundred AI use cases now, and that figure will rise, but we don't have a magic number that we aim to attain," says Prasad Ramakrishnan of Freshworks. "Rather, we will only implement those that we are fairly certain will generate value for us and our customers."

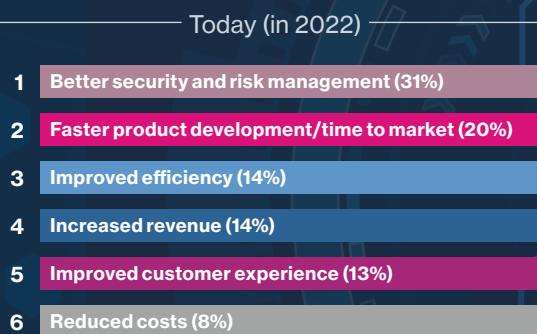


The survey respondents report solid returns from AI in a variety of areas, but the most tangible cited thus far lean toward security and risk management. Although a large number cite important AI-derived gains in faster product development and reduced time to market, relatively few executives as yet point to significant top-line returns from increased revenue.

As a group, the surveyed organizations expect to alter this picture. By 2025, net additions to revenue are expected to

By 2025, net additions to revenue are expected to be the most tangible form of return gained from AI – another sign of companies' growing ambitions for its role in their businesses.

Figure 2: Ranking of the most tangible areas of benefit from AI use today, and expected in 2025 (% of respondents)



Source: MIT Technology Review Insights survey, 2022

be most tangible form of return gained from AI – another sign of companies' growing ambitions for its role in their businesses.

Many otherwise active adopters of AI have struggled to create new revenue streams from its use. Power-engineering company Cummins started using AI five years ago to provide value-added services to its customers, such as advice to users of its engines on how to improve fuel economy or steps to take to address a parts failure. However, according to Sherry Ahholm, the company's chief digital officer, customers proved reluctant to pay additional fees for such services, instead considering the latter to be intrinsic to the product.

"This shifted our thinking on what we were doing with AI and the data generated from our engines," says Ahholm. Cummins changed the focus of its AI efforts to prognostics – predicting when certain engine parts will fail. This allows the company to suggest replacing those parts during scheduled maintenance – thus avoiding more costly warranty replacement work later. "Achieving a reduction of just 1% can be worth millions of dollars to the company," says Ahholm.

Other executives we interviewed stress that their firms are spreading their AI-related investment across many different types of use cases, and that they are generating value in a variety of ways. An example is consumer and medical products provider Johnson & Johnson. "Overall we've seen increased productivity, better risk mitigation from human error, and faster and more insight-driven decision making," says Rowena Yeo, the company's chief technology officer. Acceleration has proven to be a particularly important benefit, she adds, citing the example of an AI-powered disease forecasting model that helped the company pinpoint covid-19 hotspots and better target its clinical trials. Looking ahead, Yeo expects AI's contribution to accelerating clinical trials to directly impact revenue-generation.

Vittorio Cretella, chief information officer at Procter & Gamble (P&G), another global consumer products company, similarly credits AI for improving the company's innovation capabilities, "shortening product development time thanks to simulation and modelling, enabling more granular consumer research, and closing the loop between product innovation and consumer feedback." This, says Cretella, will ultimately translate into top-line gains.

AI use case development to 2025: Selected company examples

How will organizations be generating value from AI in 2025? The executives we interviewed shared several AI use cases they are planning to take forward in the coming months and years.

Rowena Yeo, Johnson & Johnson	Molecular modelling in drug discovery Enhanced chatbots for employee and customer interaction Accelerating clinical trials
Mike Maresca, Walgreens Boots Alliance	Micro-fulfillment centers powered by AI and robotics More precise prediction of inventory needs using analysis of omnichannel transaction data
Vittorio Cretella, Procter & Gamble	Automation of AI models (see case study, page 13) Climate change: Optimizing energy and water consumption in manufacturing
Masashi Namatame, Tokio Marine	Reducing risk in claims assessment relating to natural disasters (see case study, page 8) Refining underwriting through monitoring and analysis of driver behavior
Marc Kermisch, CNH Industrial	"The sustainable tractor": assessing the environmental footprint of tractor components (see case study, page 18)
Sherry Ahholm, Cummins	Prognostics: predicting failure of engine parts to streamline service and reduce warranty costs Improving product design and engineering
Jeremy Pee, Marks & Spencer	Expanded product personalization for omnichannel experiences Optimizing promotions and markdown
David Hogarth, Virgin Australia	Personalization of customer experience Next-gen retailing platform, including offers and dynamic pricing

Meeting the challenges of scale

For all their development of AI use cases, companies have often found the generation of benefits from them to fall short of expectations. A common refrain among technology leaders is that AI use cases have proven difficult to scale.

For S&P Global's Swamy Kocherlakota, this remains the principal challenge to surmount in the years to come. "We're spending a lot of time trying to work out how we

can apply our AI, machine learning, and NLP models at scale," he says.

The survey respondents cite internal rigidity – of organizational structures and of processes – along with budget constraints for new technologies as likely impediments to their plans for expanding and scaling AI use cases (respondents in North America are particularly

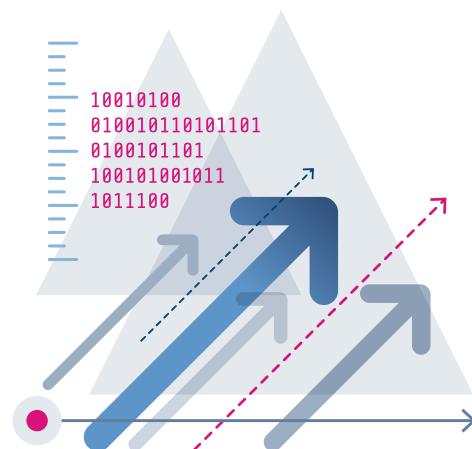
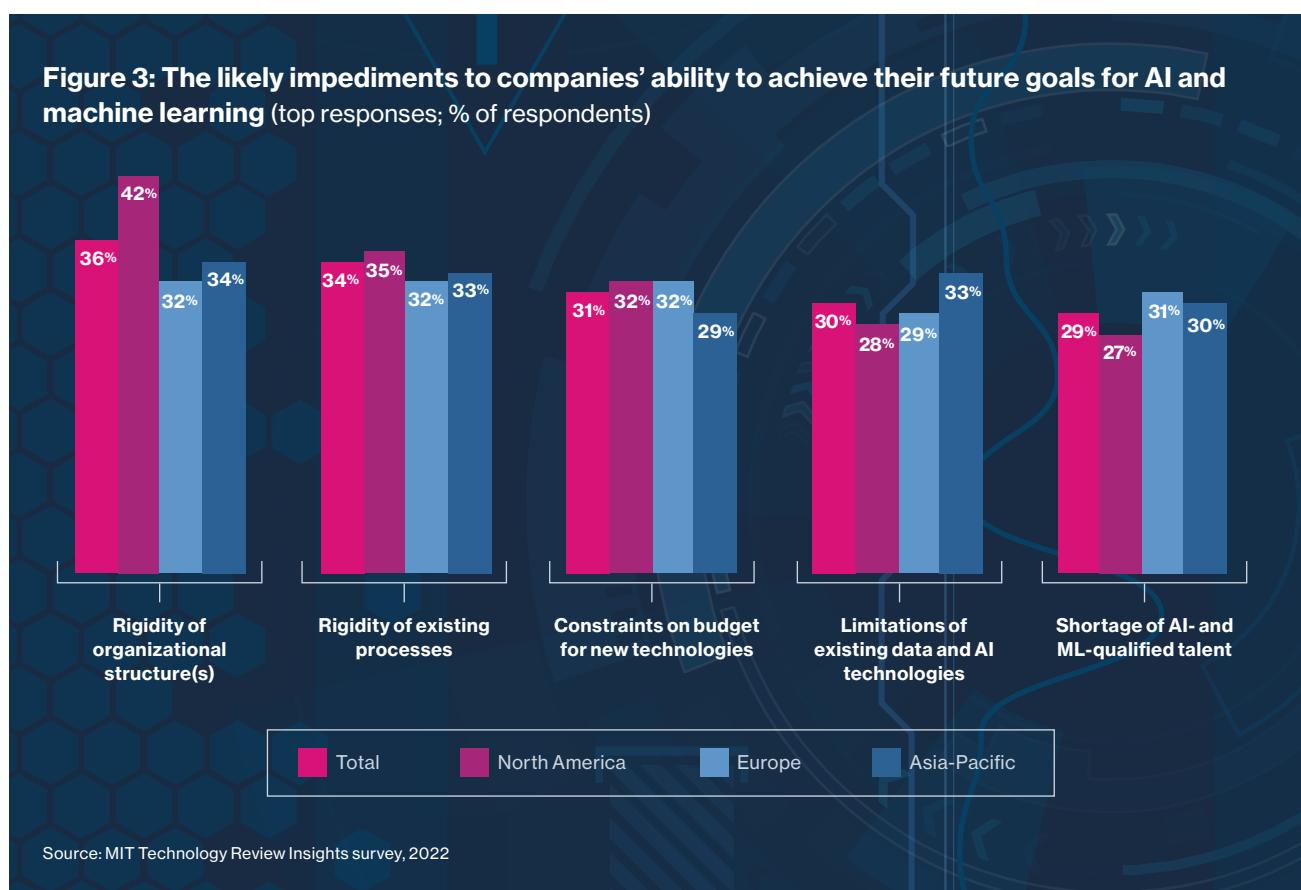


Figure 3: The likely impediments to companies' ability to achieve their future goals for AI and machine learning (top responses; % of respondents)



concerned about being held back by structural rigidities). Other constraints also loom large: limitations of existing data and AI technologies (cited most frequently by Asia-Pacific respondents) and shortages of AI-qualified talent (see Figure 3, page 11).

Elsewhere, the respondents emphasize the data challenges they face in the endeavor to embed AI more firmly in their business: 72% say that problems with data are more likely than other factors to jeopardize the achievement of their AI goals between now and 2025.

Mike Maresca, global chief technology officer at pharmaceutical retailer Walgreens Boots Alliance, places data at the top of the list of AI challenges the company needs to address, even after having upgraded its data infrastructure. “We now have the right data platform, the right quality tools, and the right governance in place,” he says. “But ensuring the data quality remains high, while we enhance our algorithms over time to continue driving right business outcomes, is a key challenge as we scale.”

That view is echoed by Rowena Yeo of Johnson & Johnson: “Data is one of the biggest challenges we face in [scaling AI], all the way from data acquisition to ingesting data, to managing it, and to ensuring the quality of the data.”

Such concerns help explain why the survey respondents believe that, after development of their AI talent and skills base, their future investments in improving data quality will be the most instrumental of a series of steps taken to progress their AI use cases.

Respondents emphasize the data challenges they face in the endeavor to embed AI more firmly in their business: 72% say that problems with data are more likely than other factors to jeopardize the achievement of their AI goals between now and 2025.

Figure 4: The most instrumental investments in helping companies to generate benefits from AI (% of respondents)

■ 2020 to 2022 ■ 2022 to 2025

Talent and skills development

26%
41%

Improving data quality

24%
29%

Machine learning infrastructure

28%
28%

Business intelligence/analytics infrastructure and tools

32%
27%

Data governance and management

29%
24%

Open-source standards and formats

22%
24%

Democratizing data (internal and/or external data sharing)

32%
23%

Unified data infrastructure

24%
23%

Process re-engineering

20%
22%

Cloud migration

19%
22%

Acquisition of AI start-up(s)

19%
20%

Data acquisition

24%
16%

Source: MIT Technology Review Insights survey, 2022

For Jeremy Pee of Marks & Spencer, the data challenges of scaling extend further: “One part of the challenge is building the infrastructure, building the data confidence, making it searchable, findable, trusted, and well-governed. The other part is making it efficient for your data scientists to build intelligence and production scalability. How do you go from a single model to building and supporting hundreds of them? If you don’t solve this part, you just end up creating a lot of inefficiency and frustration. The result – trust starts to break down.”

Next, we examine how companies plan to solidify the data foundations for the next phases of their AI development.

“Enabling the ‘democratization’ of AI involves building a set of algorithmic platforms that have intuitive front ends.”

Vittorio Cretella
Chief Information Officer, P&G

Procter & Gamble (P&G): Automating to scale

How does a multinational company that has already implemented a few hundred AI use cases develop and scale over a thousand of them? For P&G, the answer lies in automation. “We aim to develop more and more AI use cases over the next couple of years,” says Vittorio Cretella. “To do that, we need to automate the entire AI lifecycle, including data integration, model development, and model maintenance.” Amid steady spending levels across most of P&G’s data and IT operations, automating AI will be a focus of new investment in the coming months and years, he says.

In Cretella’s vision, AI automation will enable greater use-case scale in multiple ways. One is through the building of automation toolkits and workbenches that will boost the efficiency of model building and management. Machine-learning suites available from cloud hyper-scalers (such as Microsoft, Google and Amazon) will help, but P&G’s approach will also involve plugging in solutions from start-ups and open-source platforms as well as some developed internally, says Cretella. “Automation will allow us to deliver more models with consistent quality while effectively managing bias and risk,” he says.

P&G’s automation strategy also aims to unleash the human dimension of use-case development, says Cretella. He explains: “We have around 200 data scientists working on use cases in each of our

business lines, but that’s not enough to build the scale that we need. We need to allow a larger group of employees to configure key algorithms.” Enabling such “democratization” of AI involves building a set of algorithmic platforms that have intuitive front ends. “This means that business analysts can define the algorithm parameters and choose the features,” says Cretella. “We don’t need data scientists to do the coding.”

The company has a successful model to guide the development of such platforms. This is its recently patented, centralized “neighborhood analytics” platform, which Cretella describes as a “complex, multi-model algorithmic solution that clusters stores and neighborhoods based on consumer demographic and on demand and consumption signals.” It is used for a multitude of sales, marketing, distribution, and other use cases in each individual store area, says Cretella. Data scientists and analysts in every part of the business have access to the platform. Building modelling capabilities on top of it that allow non-scientists to configure and run standard AI models at scale has furthered democratization.

The new platforms to come will expand such access, says Cretella. “By lowering the entry barriers to employees to engage in model development,” it will also help P&G meet its ambitious goals for expanding AI use cases and generating greater value from them.

05 The data priorities

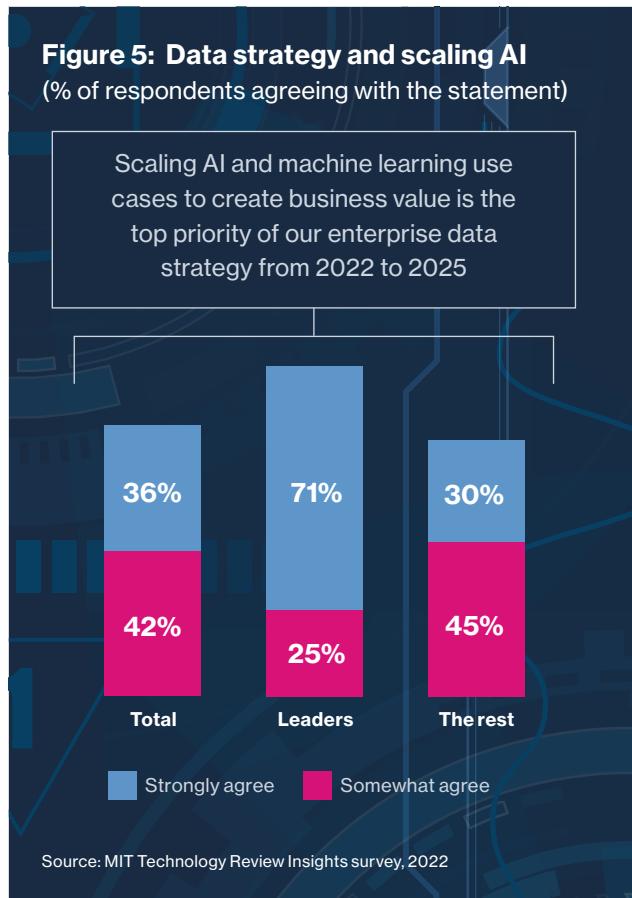


IOs and other technology leaders clearly understand the necessity of putting their data management and infrastructure in good order to be able to realize their AI ambitions. Over three-quarters (78%) of the enterprise technology leaders we surveyed say that scaling AI and machine learning use cases to create business value is the top priority of their enterprise data strategy over the

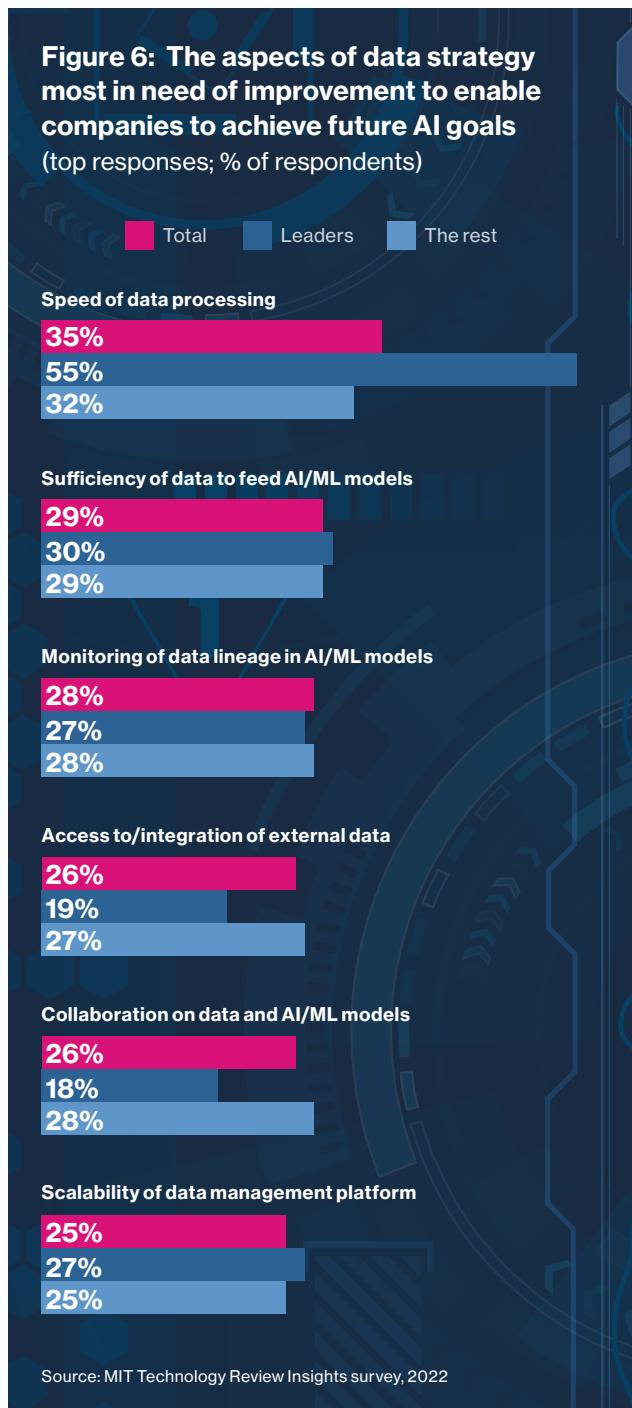
next three years. The AI leaders are almost unanimous on this point, 96% agreeing with the statement.

“Companies have so much more to achieve from an AI perspective,” says Prasad Ramakrishnan of Freshworks. “Challenge number one is to invest to build a good, solid, scalable data architecture and computing architecture, bringing more processing power to bear to handle extremely large volumes of data.”

The survey respondents concur. Asked which aspects of their company’s data strategy is most in need of improvement in order to support their AI goals, speed of data processing is at the top of their list. This is followed in order of importance by four pressing needs: ensuring the sufficiency of data to feed AI models; improving the monitoring of data lineage in those models; improving the organization’s access to, and integration of, external data;



Over three-quarters of the enterprise technology leaders we surveyed say that scaling AI and machine learning use cases to create business value is the top priority of their enterprise data strategy over the next three years.



and enabling greater collaboration on development data and AI models.

The leaders are categorical about their need to boost data-processing speeds: 55%, compared with 32% of the rest, deem this critical to realizing their AI objectives.

At least some of the platform investment made by companies will go toward unifying their data platform for analytics and AI. Over two-thirds of respondents – and nearly all leaders – say this is crucial to the success of their enterprise data strategy.

To address these and other data management challenges, respondents' companies are planning to significantly boost investment in upgrading their data capabilities between now and 2025. The investment intentions are particularly aggressive among the leaders, providing perhaps the clearest indication of how this group aims to attain AI-driven status within the next three years.

Improving data security is all respondents' top investment priority. Leaders, however, will increase their spending in this area by an average of 101% over the next three years, nearly double that of the rest of the sample. The differentials are similar for other investment targets. The leader group will invest 85% more over this period on data governance, 69% more on new data and AI platforms, and 63% more on existing platforms. The analogous figures for the rest of the survey sample are 46%, 35%, and 38%, respectively (see Figure 7, next page).

At least some of the aforementioned platform investment by companies will go toward unifying their data platform for analytics and AI. Over two-thirds of respondents (68%) – and nearly all leaders (99%) – say this is crucial to the success of their enterprise data strategy (see Figure 8, page 17).

Priorities in focus

All of the technology leaders we interviewed for the study have overseen modernization of their data platforms and architecture to one degree or another in recent years.

While not all intend significant new investments in the immediate future, most emphasize that modernization is a work in progress, and they will be acting accordingly to ensure their specific data objectives are met.

For air carrier Virgin Australia, the priorities revolve around improving data management. In the past 12 months, it has implemented a new, cloud-based data technology platform and a customer data platform to support its personalization objectives. That has also involved building an entirely new data-engineering team. The ongoing challenge, says David Hogarth, the company's chief information officer, is bringing together the disparate data sets the airline uses into one model. "That will put us in a good position to build the long list of AI use cases that our business users are demanding from us," he says.

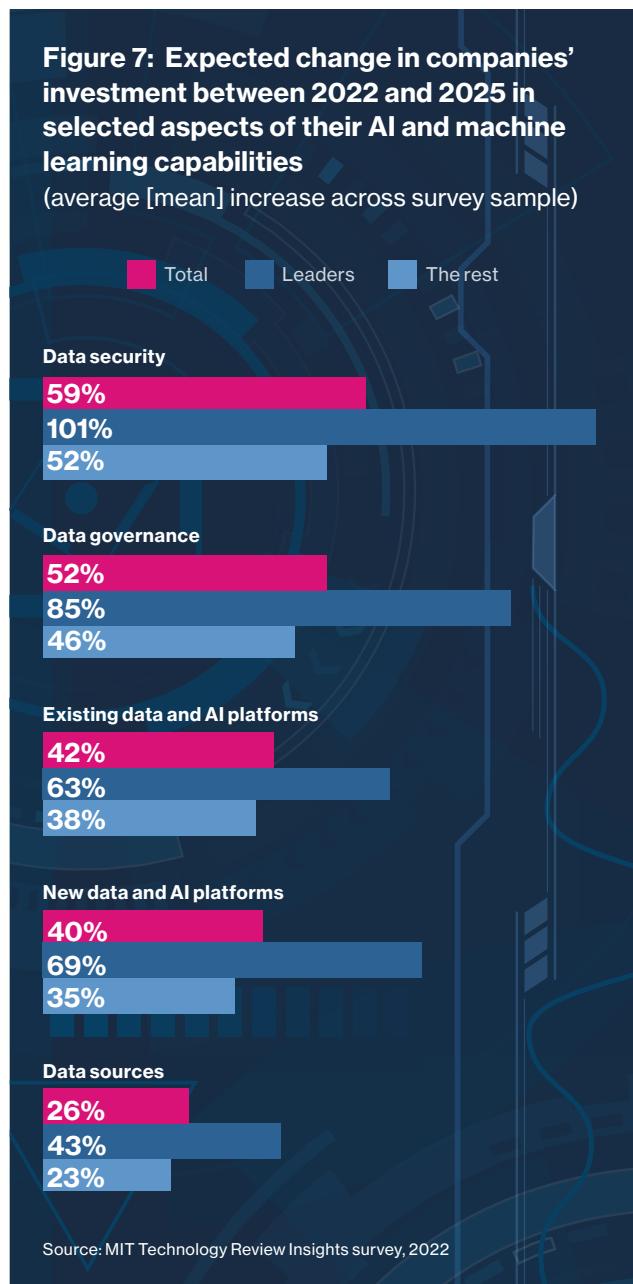
With the technology platform it now has, adds Hogarth, properly unifying the data will also enable the company to

"Many companies don't really know what they have in terms of their entire data estate or how they measure quality across it. We're addressing this in our new platform by putting cataloging, searchability, data quality management, and other capabilities around every data element. We need to do all of this at the speed necessary to respond to customer and business needs."

Jeremy Pee
Chief Digital and Data Officer,
Marks & Spencer

democratize its use safely and securely. "Our users don't see value in acquiring the data themselves. Providing them with high-quality, consolidated insights from the data allows them to rapidly conduct analytics and achieve value quickly. That's a highly attractive proposition to them."

Continuing to build scale and speed of data processing is a top priority for engineering company CNH Industrial, according to its chief information officer Marc Kermisch. "A lot of the modernization we are engaged in is transforming databases and warehouses to be able to manage both unstructured and structured data in the



cloud at scale and at speed,” he says. “We aim to develop event-based capabilities that allow us to do modeling and analysis at or close to real time.”

At Marks & Spencer, says Jeremy Pee, the focus is on augmenting the data platform it has recently built with additional capabilities. “What’s harder to do than build models and pipelines is solve for quality, use, and understanding of data. Many companies don’t really know what they have in terms of their entire data estate or how they measure quality across it. We’re addressing this in our new platform by putting cataloging, searchability, data quality management, and other capabilities around every data element. We need to do all of this at the speed necessary to respond to customer and business needs.”

Multi-cloud and open

The development of organizations’ AI and machine-learning capabilities is facilitated by a multi-cloud strategy – the use of services from two or more public cloud vendors. A large majority of the survey respondents (72% overall, and 92% of leaders) believe a multi-cloud approach ensures the most flexible possible foundation for AI development. Among other advantages, it enables AI teams to choose the most suitable platforms for the development of different use cases having specific resource requirements. The latter include the sourcing, storage, or processing of data.

A multi-cloud strategy and an open approach to data architecture and standards often go hand-in-hand. “Having a multi-pronged, multi-cloud approach, and then incorporating APIs and microservices as part of our data architecture, are key for us,” says Rowena Yeo of Johnson & Johnson.

Walgreens Boots Alliance has developed its new data platform with a focus on the use of open standards and open data, says Mike Maresca. “At our core, we support the use of open source technologies on our cloud data platform, which increasingly supports integration across different cloud providers. Open source standards and the ability to integrate cloud services across providers is important to our efforts to fully embed AI and machine learning in our business.”

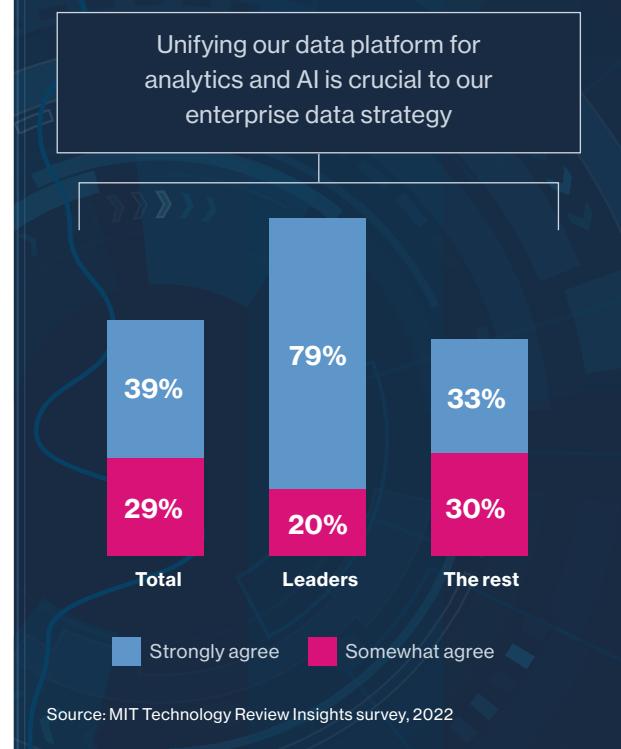
Open standards are also facilitators of API-based data sharing, which companies in the healthcare community such as those mentioned above rely on for AI-based innovation. It is increasingly becoming the case also in

other industries, such as manufacturing (see “CNH Industrial: AI, open data, and the sustainable tractor”, page 18). “Open-source standards are really important for our ability to take AI and machine learning to the next level,” says Kermisch.

“Having a multi-pronged, multi-cloud approach, and then incorporating APIs and microservices as part of our data architecture, are key for us.”

Rowena Yeo, Chief Technology Officer & Global Vice President, Technology Services, Johnson & Johnson

Figure 8: The importance of a unified data platform (% of respondents agreeing with the statement)



CNH Industrial: AI, open data, and the sustainable tractor

The contributions that AI is making to sustainable agriculture are many and varied. Precision farming, optimizing water and chemical use, locating carbon sinks, enabling urban farming to reduce deforestation – these are just a few of the existing use cases.

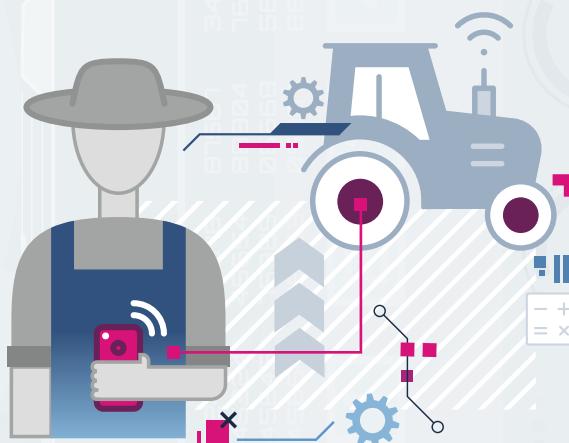
CNH Industrial plans to make another AI-enabled contribution to the cause – a sustainable tractor.

According to Marc Kermisch, the ability to access and analyze external data is vital to develop this use case. “We need to understand the environmental footprint of each component we put into the tractor,” he says. “That requires obtaining data from each manufacturer on, for example, its wastewater production, its energy consumption, and its hazardous material treatment. If a part contains hazardous materials, are they resolvable over time? Does the part need to be treated in a special way at the end of the tractor’s life to avoid damaging the environment? Is it recyclable?”

There are large supply chains that agricultural equipment producers like CNH Industrial need to work with to ensure their products leave the lightest possible footprint. To make the right decisions in creating a sustainable tractor, says Kermisch, the necessary external data must be accessible and accurate, but those conditions do not yet exist. “Some vendors may have created third-party databases, but we have no way of validating them to ensure we can trust the data,” he says. “Getting this wrong has enormous implications for sustainable farming.”

Kermisch is hopeful that the Science-Based Targets initiative (SBTi)⁴ – a not-for-profit partnership dedicated to fostering sustainability metrics – can help his industry meet this challenge. A good solution, he says, would be for SBTi to anonymize the data generated by agricultural equipment producers, resulting in public data sets that can easily be validated. That would ease concerns about sharing commercially sensitive data. Such solutions exist in other fields of engineering, he says, and one should be possible for agricultural machinery.

With widely available, accurate, and reliable data, the AI-enabled tractor will enable farming that is not just more efficient but also less carbon-intensive.



Precision farming, optimizing water and chemical use, locating carbon sinks, enabling urban farming to reduce deforestation – these are just a few of the existing use cases for AI in sustainable agriculture. CNH Industrial plans to make another: a sustainable tractor.

An industry lens on data and AI

Among this study's objectives have been to compare the AI ambitions of different industries and to determine the extent to which industry approaches to removing impediments to AI development vary. **Three areas of difference stand out:**

1 Retailers and manufacturers are more ambitious about becoming AI-driven.

Within the leader group – where AI is expected to be critical in at least five functions by 2025 – respondent companies from the retail/consumer goods and automotive/manufacturing sectors are the most prominent. Life science and healthcare organizations are also well-represented (see Figure 9).

2 Data constraints on AI development loom especially large in three industries.

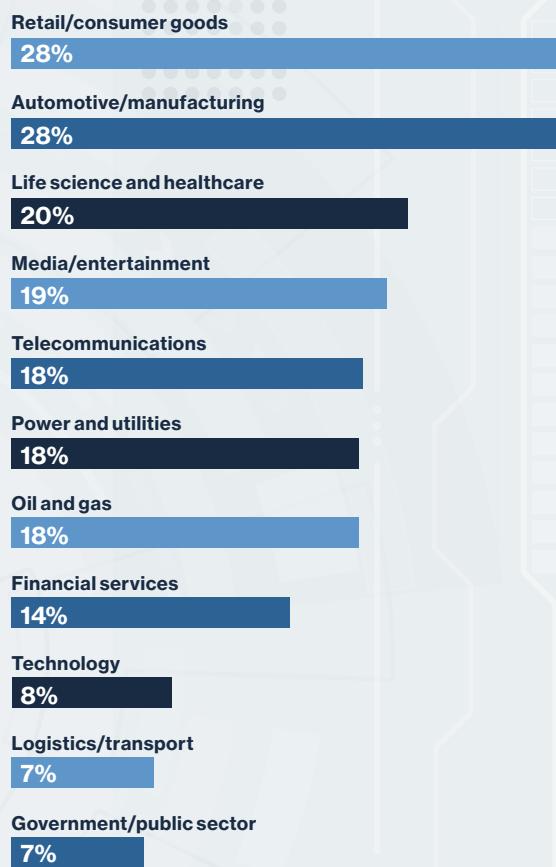
Among real estate and construction respondents, almost double the percentage of those in the overall sample (58% vs 30%) cite limitations of existing data technologies as an impediment to AI development. Some 40% of automotive/manufacturing respondents say the same. And 82% of executives in the latter sector (compared with 72% overall) state that data-related problems are more likely than other factors to jeopardize achievement of their future AI goals. That belief is shared by 80% of financial sector executives and 78% of those in real estate and construction firms.

3 Financial service providers display the strongest investment growth intentions.

Expected spending growth on data capabilities in the financial sector dwarfs that in others across several line items. For example, investment to bolster data governance will increase by 74% between now and 2025, according to financial industry respondents, compared with 52% for the sample as a whole. The

analogous differential relating to existing data and AI platforms is 61% vs 42%, and for new platforms – 58% vs 40%. Spending growth in these areas by retail/consumer goods and automotive/manufacturing firms will also exceed the sample average.

Figure 9: Where AI leaders are most numerous (% weight of leaders within each surveyed industry)



Source: MIT Technology Review Insights survey, 2022

Retailers and manufacturers are more ambitious about becoming AI-driven, but financial service providers display the strongest investment growth intentions.

06 Conclusion

As much as organizations have done to develop AI capabilities and embed them in the business, their journey to becoming AI-driven has only just begun. CIOs recognize that their organizations have thus far only scratched the surface of the efficiency, speed, innovation, and other gains that the use of AI and machine learning can generate across different functions. They also recognize that the data, talent, and other foundations they are putting in place to support AI development cannot remain static. The foundations must evolve not just to enable the critical scale of use cases to be reached, but also to keep pace with future advances in the science of AI and the demands they may pose for additional power, expertise, and process change.

CIOs cannot expect the technology foundations they are putting in place today for AI to be future-proof. However, this research points to a handful of attributes that technology leaders can instill in their data and other technology foundations in order to facilitate a smoother evolution.

Democratization. The greater the number of employees in an organization who can configure and improve AI algorithms, the more AI-based innovations are likely to

materialize. Many CIOs are looking to “citizen data scientists” – data-literate employees without specialist data science training – to rise to this challenge. Democratization is all the more important while qualified AI specialists and data scientists remain in limited supply. The infrastructure modernization that CIOs pursue should aim to widen employee access to data needed for algorithm development.

Openness. Few fields of technology have benefitted as much from open, multi-partner collaboration as AI. CIOs know that their companies’ future success in innovating with AI will rely at least in part on the data, insights, and tools they are able to source externally. Data technology that favors open standards and open data formats is well placed to facilitate such collaboration.

Multi-cloud. The multi-cloud approach that technology leaders favor to help scale AI in their business can be challenging to manage, due to the complexities involved in monitoring and optimizing AI projects across several vendor environments. However, platforms with centralized capabilities (MLOps, for example) are increasingly an option to manage the complexities. And it’s hard to argue with the access multi-cloud affords to data processing power on-demand and new, cloud-based AI solutions.

CIOs recognize that the foundations they are putting in place to support AI development must evolve to keep pace with future advances in the science of AI and the demands they may pose for additional power, expertise, and process change.

About MIT Technology Review Insights

MIT Technology Review Insights is the custom publishing division of MIT Technology Review, the world's longest-running technology magazine, backed by the world's foremost technology institution – producing live events and research on the leading technology and business challenges of the day. Insights conducts qualitative and quantitative research and analysis in the US and abroad and publishes a wide variety of content, including articles, reports, infographics, videos, and podcasts. And through its growing MIT Technology Review Global Insights Panel, Insights has unparalleled access to senior-level executives, innovators, and entrepreneurs worldwide for surveys and in-depth interviews.

From the sponsor

Databricks is the data platform company that unifies all your data, analytics and AI on one platform. More than 7,000 organizations worldwide – including Comcast, Condé Nast, H&M and over 50% of the Fortune 500 – rely on the Databricks Lakehouse Platform to unify their data, analytics, and AI. Databricks is headquartered in San Francisco, with offices around the globe. Founded by the original creators of Apache Spark™, Delta Lake and MLflow, Databricks is on a mission to help data teams solve the world's toughest problems. To learn more, follow Databricks on [Twitter](#), [LinkedIn](#) and [Facebook](#).



databricks

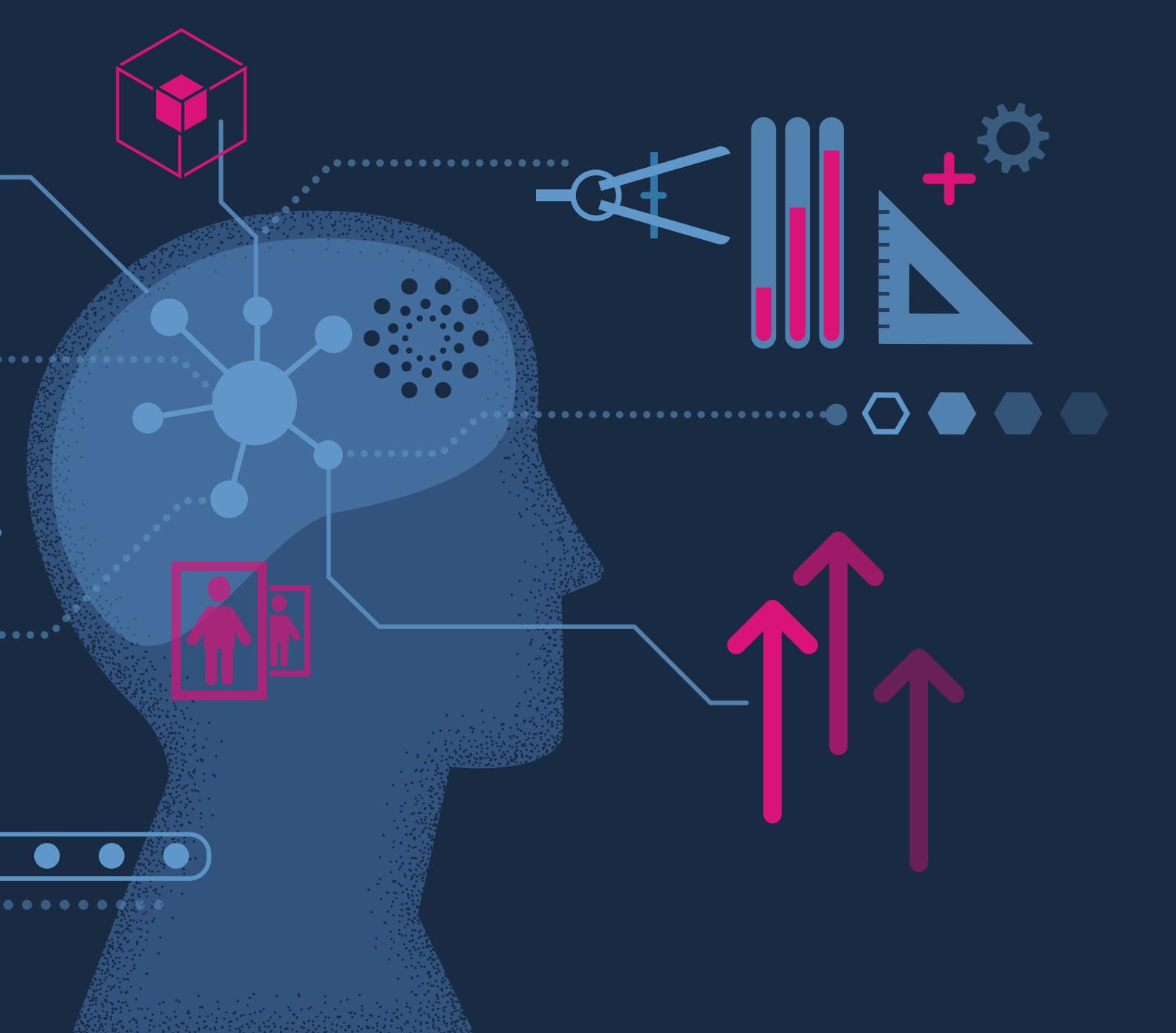
Footnotes

1. The leaders comprise 14% of the overall survey sample.
2. For a discussion of the organizational and cultural barriers impeding companies' progress in embedding AI, see "Building the AI-Powered Organization," Harvard Business Review, July-August 2019. The talent shortages constraining wider adoption of AI are discussed in "The AI Talent Shortage Isn't Over Yet," The Wall Street Journal (content by Deloitte), October 16, 2020.
3. AI and machine learning are not one and the same, the latter being but one field in the wider science of AI. (Natural language processing and neural networks are other constituent fields.) For ease of reference, however, we will use the singular term of AI throughout this study, except where the discussion specifically focuses on machine learning.
4. The SBTi was formed by CDP, a charitable organization; the United Nations Global Compact; the World Resources Institute (WRI); and the World Wide Fund for Nature (WWF).

Illustrations

Cover and interior artwork assembled by SSD with elements from Adobe stock at stock.adobe.com.

While every effort has been taken to verify the accuracy of this information, MIT Technology Review Insights cannot accept any responsibility or liability for reliance by any person in this report or any of the information, opinions, or conclusions set out in this report.



MIT Technology Review Insights

-  www.technologyreview.com
-  [@techreview](https://twitter.com/techreview) [@mit_insights](https://twitter.com/mit_insights)
-  insights@technologyreview.com