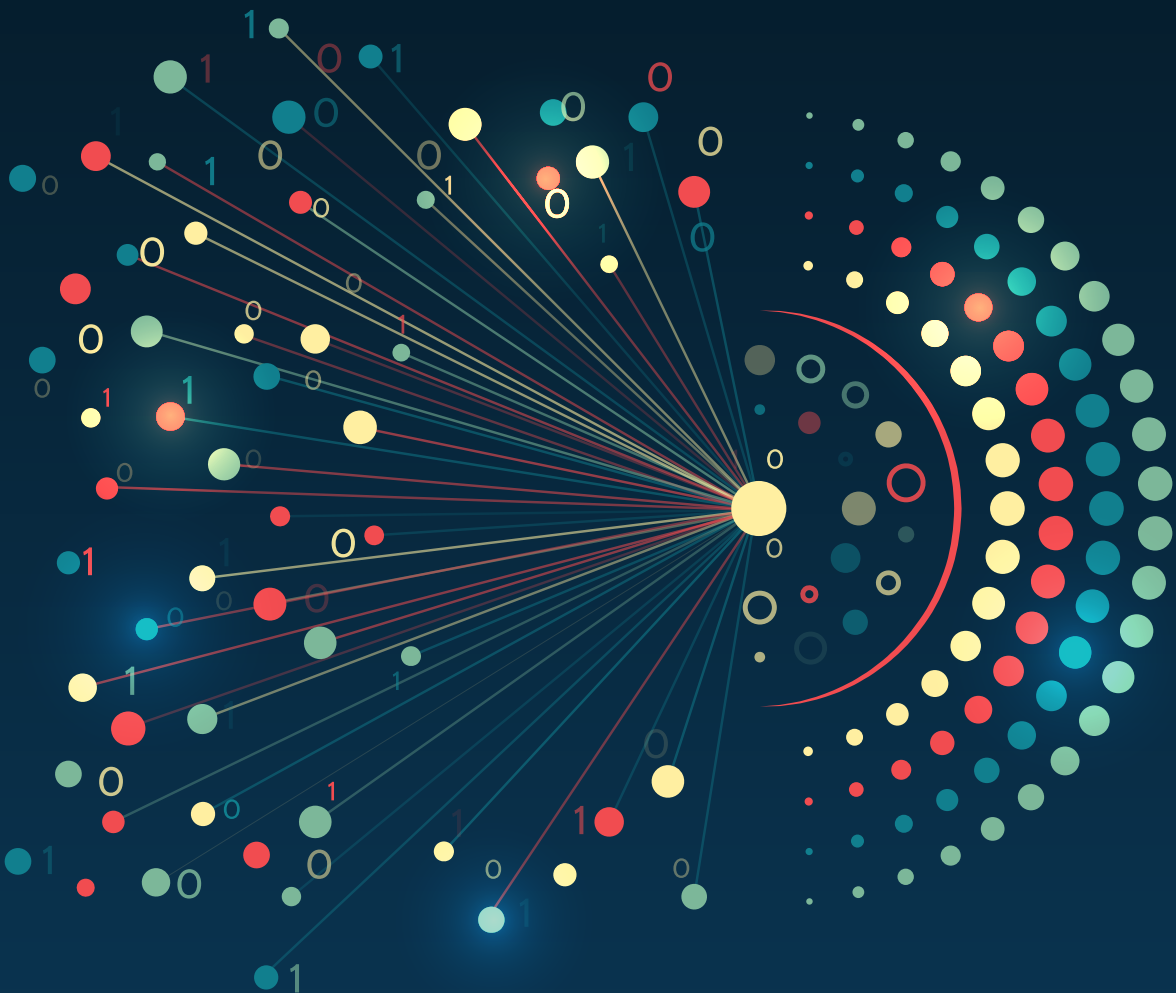
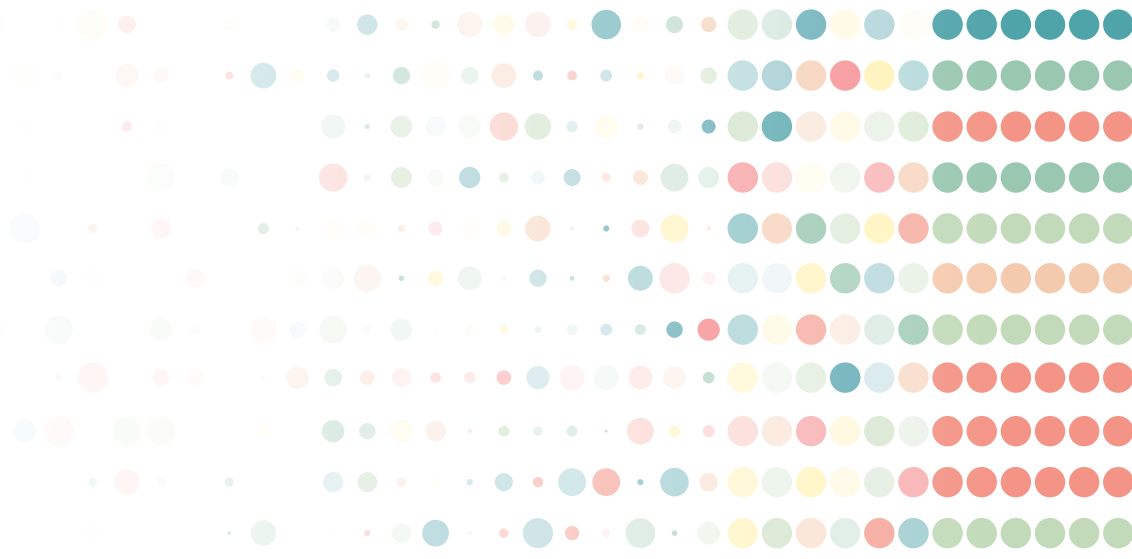


Scaling AI agents with trustworthy data



“Scaling AI agents with trustworthy data” is an MIT Technology Review Insights report sponsored by Google Cloud. This report, which is based on survey research and executive interviews, seeks to understand how well organizations’ data systems support their agentic AI use cases. Denis McCauley was the author of the report, Virginia Wilson was the editor, and Nicola Crepaldi was the publisher. The research is editorially independent and the views expressed are those of MIT Technology Review Insights.

Ashutosh Mishra, Vice President, Data Strategy and Platform, Deutsche Telekom



CONTENTS

- 01 Executive summary 5
- 02 The value of AI agents today 6
 - Use cases and objectives..... 7
- 03 Overcoming the legacy burden..... 9
 - Addressing legacy systems head-on at Shopify..... 11
 - Bringing dark data to light..... 11
- 04 Wanted: Integration, access, velocity, and context..... 12
 - Fewer silos, greater access..... 12
 - Real time and relevant..... 13
 - Moving from data to knowledge at Deutsche Telekom..... 13
- 05 Building agentic readiness 14
 - HCA Healthcare’s journey to an AI-native data estate 15
 - Building for scale 15
- 06 Conclusion: A framework for success 16

Foreword

AI is reengineering the global economy. In the agentic era, businesses are being supercharged by autonomous agents that perceive, reason, and act to unleash human potential. This introduces three major shifts that flip the script on how we think about data powering AI.

First, we are moving from human scale to agent scale. You will face orders of magnitude more workload as agents monitor and operate your business around the clock at digital speed. Your data platform now needs to support always-on, high-velocity, and autonomous operations.

Second is a shift from reactive intelligence to proactive action. Previously, we built reactive architectures to tell us what happened yesterday. Now, your agents must act on your behalf, and do so accurately and in real time.

Third is a shift from data to context. It is no longer enough just to discover and query your data. You must elevate this data with a knowledge flywheel that understands the relationship between data assets, semantics, and how the data is commonly used. Without a deep understanding of the full data estate, including unstructured dark data, you do not have the quality and trust to activate agents at scale.

Many vendors talk about their data platforms as a System of Intelligence. But those are the last generation of data platforms. The goal is no longer just to know. The goal is to act. To do this, you need a System of Action.

The so-called “modern” data stacks that are glued together with disjointed parts were not built for the agentic era. When we analyze these legacy architectures under agentic

load, we see they break. They create walled gardens that fracture access controls and dissolve governance perimeters. They suffer from a trust gap, where agents act on raw data without business understanding, leading to hallucinations. They break the reasoning loop by separating the thinking from the doing, meaning agents miss the window to act in real time. Finally, because they are a patchwork of rented models and infrastructure, they lead to a cost spiral. At agent scale, your AI should be your greatest asset, not an unpredictable financial liability.

You need a new architecture: an Agentic Data Cloud. This new category of architecture evolves the enterprise data platform from a static, fragmented repository into a dynamic, unified knowledge platform. To survive agent scale, it must be AI-native from the hardware to the software. It must be flexible, tearing down walled gardens to activate your data wherever it lives. Above all, it must be trusted. Deep business context and strong governance must provide active oversight, ensuring every automated action is permissioned, explainable, and compliant.

At Google Cloud, we have rearchitected our entire data portfolio to natively support these autonomous workflows. By integrating AI models directly with Google Cloud’s differentiated data infrastructure, we eliminate the complexities of scale, governance, and cost, helping enterprises bridge the gap between intelligence and trusted action.

Andi Gutmans, Vice President and General Manager, Data Cloud, Google Cloud



Executive summary

Business and technology leaders need no convincing that the time of agentic AI is here. Organizations are rapidly adopting agents,¹ and few executives doubt the technology's potential to transform work. But many organizations find that realizing the desired return on investment (ROI) from AI hinges on having the right foundation, with inadequate infrastructure and data being major blockers.²

Agentic AI places considerable new demands on enterprise data systems. The shift from answering questions to taking actions means AI agents need data from across the enterprise, in all its structured and unstructured forms, and with the right business context. To make decisions and act in real time, agents also need frictionless access to the organization's operational systems – for example, those storing its supply chain, point-of-sale, or human resources data. Legacy data systems, even those updated just a few years ago, struggle to meet these demands.

As AI agents become embedded more widely in enterprise operations, the need to overcome the restrictions of legacy data systems grows more urgent. If Gartner's prediction that AI agents will augment or automate 50% of business decisions by 2027 proves correct,³ organizations must eliminate bottlenecks or risk depriving agents of the data they need to make the right decisions at speed.

This report, based on a survey of 300 data and technology executives, explores how legacy systems are limiting the effectiveness of AI agents in many organizations. It finds that a handful of organizations – the data leaders – are having greater success with

agentic AI and experiencing fewer data limitations as a result of legacy systems. These leaders offer a guide to creating the right data environment for agents to flourish and trusted systems to scale.

Key findings from the report include:

- **Few companies currently provide agentic AI with ample access to enterprise data.** Across all the surveyed organizations, AI only has access to an average of 45% of company data. That number falls to 30% or less in organizations categorized as “data laggards”. A select group, however, ensures access to over 70% of their data. These “data leaders” are having greater success with their agents than the rest.
- **Trust in agent decisions is a reflection of data readiness.** Today, only around half of surveyed organizations trust that the decisions their AI agents make are accurate and relevant. By contrast, 100% of the data leaders trust their agents' decisions, a strong indicator that reliable AI requires a reliable data foundation.
- **Data leaders find it easier to achieve agent scale and speed.** Two-thirds of data laggards say legacy data systems limit AI agent scaling (66%) and prevent agents from making decisions at speed (68%). Having largely overcome legacy data constraints, the leaders have mostly cleared these roadblocks, with just 8% reporting either constraint.
- **The pressure is on to make data estates agent-ready.** Within two years, 100% of respondents plan to be using agentic AI, with 69% expecting to use it widely. Without removing data system constraints, agentic AI will fail to deliver the desired speed and efficiencies it promises.
- **Data access and context are top priorities.** The most important initiative to enable scaling among all respondents is improving access to structured and unstructured data for AI agents. Also high on the list is enhancing data and AI governance with business context. Data leaders are also focusing heavily on the automation of data management.

02 The value of AI agents today

Moving beyond simple analysis and prediction, AI agents are capable of independent reasoning and autonomous decision-making. They also take action, executing individual tasks and organizing entire workflows.

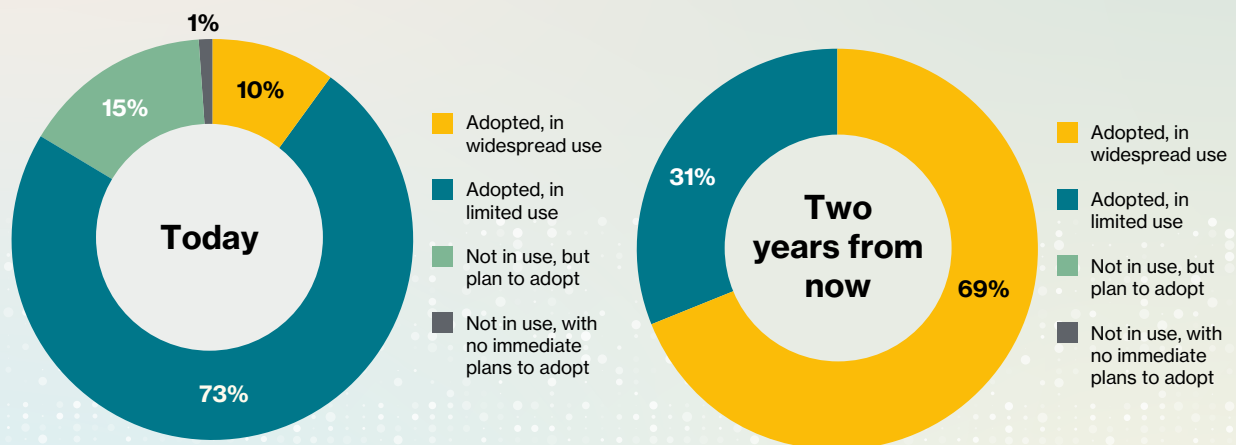
Nearly all the organizations in our survey (98%) are either already using agentic AI or plan to (see Figure 1). One in 10 (10%) are using it widely and 73% have deployed it in a more limited fashion. Another 15% of organizations plan to adopt it in the future. High-tech companies are, perhaps unsurprisingly, the most active users of AI agents today, with 20% saying they already employ them widely. Some 14% of healthcare

and life sciences organizations are also putting agents to widespread use.

Within two years, 100% of companies say they will use agents, with 69% saying they will be doing so widely, indicating a trend towards the rapid scaling of agentic AI (see Figure 1).

It is easy to understand why companies are eager to put AI agents to work. Used effectively, agents promise significant increases in team productivity and efficiency as they automate repetitive tasks and assume the management of software and infrastructure workflows, so employees can focus on more strategic work.

Figure 1: 83% of organizations are using AI agents to some degree today; all intend to use them within two years, with seven in 10 planning widespread adoption



Source: MIT Technology Review Insights survey, 2026. Figures may not total 100% due to rounding.

AI agents can adapt to changing conditions, increasing the relevance of their decisions. And they can make those decisions at high speed.

Use cases and objectives

There are three main areas where businesses are most likely to develop agentic AI use cases today:

- **Customer service:** 62% of the surveyed organizations are using AI agents to help manage customer service workflows, such as the routing of requests and issue resolution.
- **IT systems management:** 52% have deployed agents in IT systems management, where they help to manage user access and incident response, among other processes.
- **IT security:** 51% use agents to support IT security, where they are executing tasks in areas like anomaly detection and threat scanning.

Software management (such as application development and engineering) and sales and marketing are other common areas of deployment (see Figure 2).

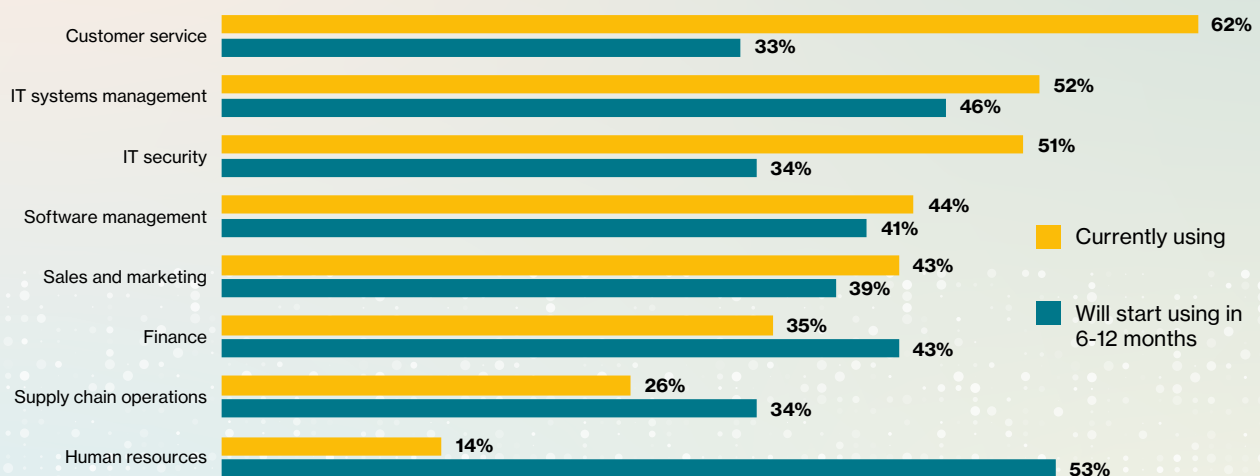
Other use cases will start coming to the fore six to 12 months from now:

- **Human resources:** over half (53%) of surveyed organizations will develop agent use cases for HR processes, such as onboarding and benefits management.
- **Finance:** 43% will put agents to work in finance areas like invoice processing and expense reporting.
- **Supply chain:** 34% will start using agents to execute tasks involving inventory management and freight routing.

“We must allow agent access to data in safe and trustworthy ways so people can make maximal use of the data, knowing that it’s completely reliable.”

Rajpreet Bajwa, Vice President,
Engineering, Data and AI Infrastructure,
Shopify

Figure 2: Organizations are most commonly using AI agents in customer service and IT today, with plans to extend them to HR and finance within a year



Source: MIT Technology Review Insights survey, 2026

“Adopting agentic AI shouldn’t be a challenge for any enterprise. Being agent-ready is the bigger challenge. A lot of enterprises really struggle with agentic AI because their data is not in a form that can be accessed.”

Ashutosh Mishra, Vice President, Data Strategy and Platform, Deutsche Telekom

Despite being early days, agentic AI is already delivering gains in many of these key areas. Nearly half of respondents (46%) say their teams are having success with AI agents in customer service. And 42% say the same for IT systems management and IT security (see Figure 3). “We’re very excited about the productivity improvement that agentic AI has enabled for us,” says Rajpreet Bajwa, vice president of engineering and data and AI infrastructure at Shopify. “We’re able to work with data faster – for example, faster programming, unit testing, and analysis.”

We have identified a select group of 25 organizations (of the surveyed 300) that are enjoying greater success with their agents than others. We call this group data leaders – those with AI systems that can access more than 70% of enterprise data. (See also “Bringing dark data to light” in Section 3.) Their superior results are evident here: nearly two-thirds of these leaders (64%)

are seeing gains from agent use in customer service and software management, and 44% in IT systems management, IT security, and sales and marketing. Data laggards (where AI can access 30% or less of enterprise data), on the other hand, report far less success in those same functions.

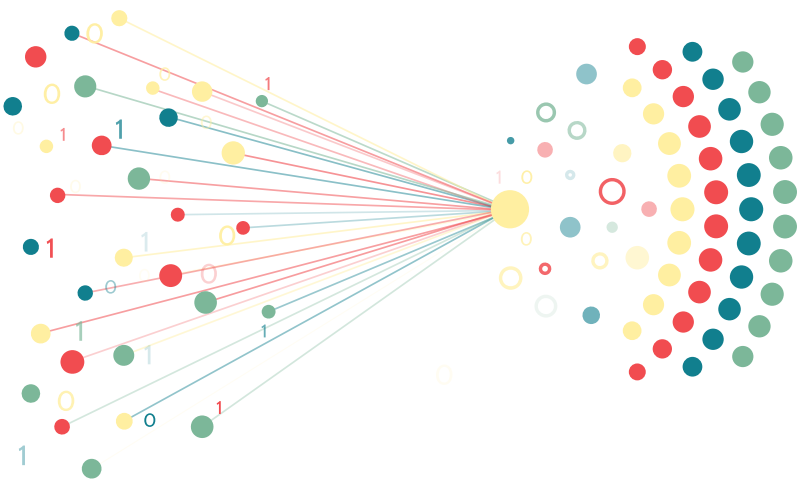
If companies scale agentic AI as planned, it could transform business and technology operations and deliver significant gains within a few years. However, like any enterprise technology advancement, realizing this potential requires overcoming cultural and process-related hurdles like legacy data systems. “Adopting agentic AI shouldn’t be a challenge for any enterprise. Being agent-ready is the bigger challenge,” says Ashutosh Mishra, vice president of data strategy and platform at Deutsche Telekom. “A lot of enterprises really struggle with agentic AI because their data is not in a form that can be accessed.”

Figure 3: The areas where organizations are having the most success with AI agents



Source: MIT Technology Review Insights survey, 2026

03 Overcoming the legacy burden



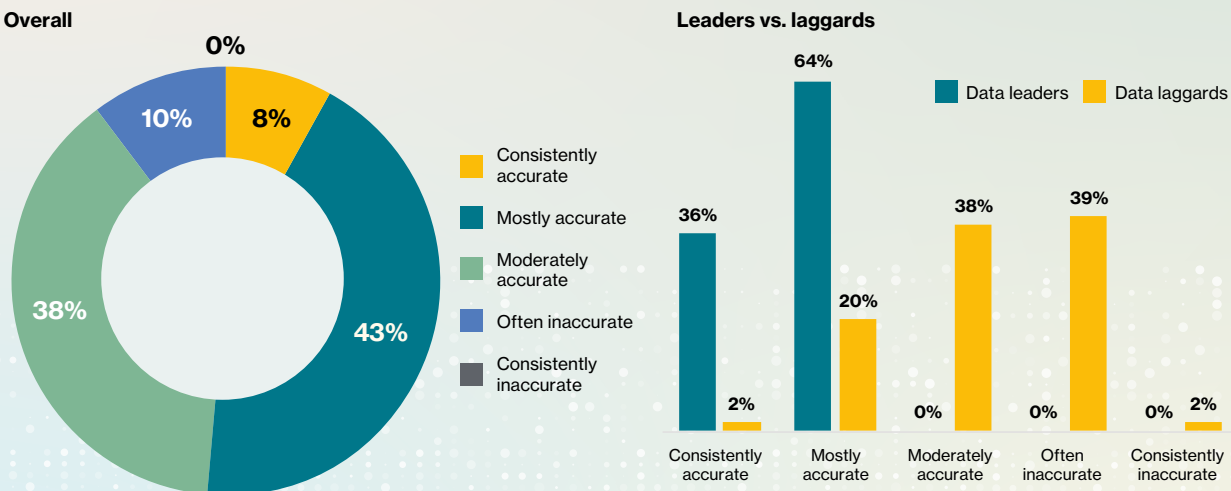
To function effectively, AI agents and agentic applications need data that is:

- 1. Multimodal**, taking not just structured form, as in traditional databases where data is organized into rows and columns, but also images, PDFs, emails, documents, audio, video, and other unstructured forms
- 2. Context-aware**, containing information about the who, what, when, and where of the data's origin and its relationships, including who is using it and for what purpose
- 3. Instantly available**, capable of being delivered to AI agents in real time from the enterprise's various repositories, such as operational systems

Legacy data systems struggle to meet these demands, compromising AI trustworthiness. Around half of the surveyed executives overall (51%), and just one-fifth of the data laggards (22%), trust the accuracy and relevance of their AI agents' decisions (see Figure 4). By contrast, this trust gap vanishes for data leaders (few of whom struggle with legacy data platforms, as shown in Figure 5): 100% of these organizations report that their agents make mostly or consistently accurate and relevant decisions, demonstrating that trust is unlocked at the data layer.

"We must allow agent access to data in safe and trustworthy ways so people can make maximal use of the data, knowing that it's completely reliable," Bajwa says.

Figure 4: Better data access leads to higher trust in accurate and relevant AI agent decision making
How accurate and relevant an agent would be today in making decisions based on accessible data, according to respondents, using their current data architecture.



Source: MIT Technology Review Insights survey, 2026. Figures may not total 100% due to rounding.

Most organizations believe legacy data systems are causing scaling and latency issues. More than half the respondents overall (55%) and 66% of data laggards say legacy data systems are preventing them from scaling agentic AI. And 52% overall – including 68% of laggards – say legacy systems are preventing AI agents from making decisions at speed. Just 8% of data leaders report the same scaling or latency issues (see Figure 5).

These legacy data challenges make it difficult for many organizations to see the desired returns from the use of AI agents. Half the respondents overall say legacy data systems are having a significant negative impact on the ROI of their agent-driven projects. Although far fewer, 32% of data leaders also say the same. This

indicates that some legacy components can remain in place amid modernization and limit agent performance. “Legacy hugely impacts ROI,” says Mishra. “Running AI agents with legacy data systems requires a lot of procurement, infrastructure cost, and engineers to maintain and develop them. That entails massive upfront investment.”

“Teams deploying agents need to demonstrate ROI, since doing so involves various costs, such as tokens,” says Andi Gutmans, vice president and general manager of data cloud at Google Cloud. “They’re trying to figure out how best to get the desired returns, and one of the ways is making sure their data is in a more cost-effective, AI-ready, and scalable state.”

Figure 5: Legacy data platforms are preventing organizations from scaling and advancing agentic AI – and this is impacting ROI

Share of respondents who strongly or moderately agree with statements about their data systems and AI agents.

Overall

55%

Our current data platform is the primary bottleneck preventing us from scaling AI agents across the enterprise

54%

We have paused or delayed the deployment of AI agents specifically to fix foundational data issues (e.g., data silos, lack of governance and business context)

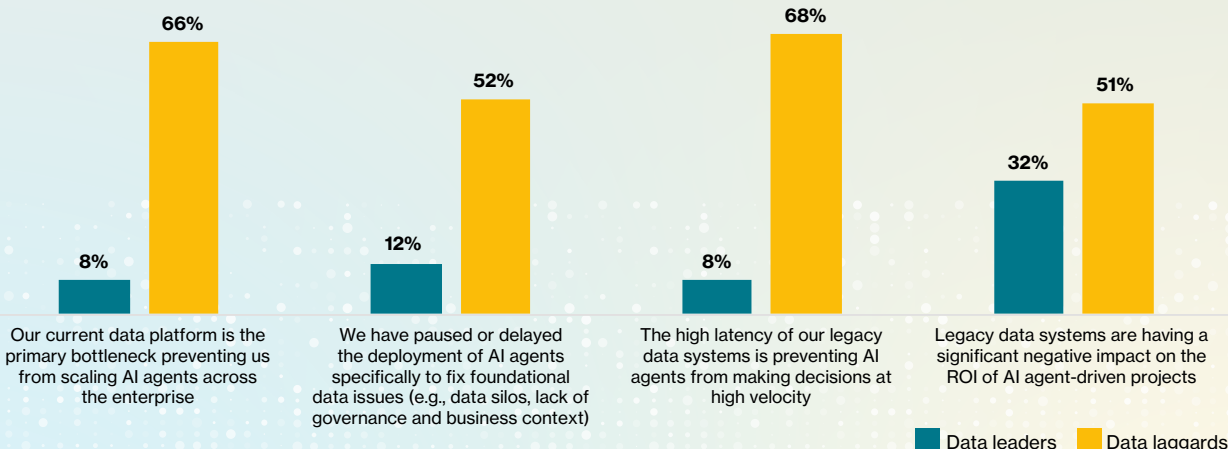
52%

The high latency of our legacy data systems is preventing AI agents from making decisions at high velocity

50%

Legacy data systems are having a significant negative impact on the ROI of AI agent-driven projects

Leaders vs. laggards



Source: MIT Technology Review Insights survey, 2026

Addressing legacy systems head-on at Shopify

When Rajpreet Bajwa first started looking after Shopify's data platform, her first order of business was to get the data foundations ready for AI agents. "We recognized how important this was, as we'd developed legacy debt over time. So we invested in new data platforms and turned off a lot of our legacy systems."

That data modernization has underpinned several successes the company has since had with its AI agents, says Bajwa. She is proud, for example, of the agentic products her team has deployed for Shopify's merchants and buyers. And the team is also continuously launching other types of agentic flows, she says. "Even more exciting are the internal

productivity gains we've made. We're now able to work with data faster, do better programming, better unit testing, and faster analysis. We're also now able to build recursive code improvement loops, where we can evaluate the output of a model that's doing analysis and then retrain it."

Bajwa's team is continuing to spend on the data foundations, she says, including reinforcing the access pillars to add more context to the company's data. "That helps turn a prototype into something that you can scale and can put a lot of business logic in," Bajwa explains. "And you can trust the agent's output when you're scaling with the right metadata and lineage."

Bringing dark data to light

Legacy data architectures are unable to address a fundamental AI constraint: "dark" and other difficult-to-access data. Dark data consists largely of unstructured data contained in images, video, PDFs, and social media posts, for example. It may also include structured data that is stored away in operational systems, like customer support logs or other old customer relationship management (CRM) data, server log files, manufacturing or logistics data transmitted by IoT devices, or old data in HR systems. "To get the most out of AI, you have to be able to activate all of your data," says Gutmans, "whether that is structured or unstructured. Data will typically be sitting in a variety of systems. It's not all going to be sitting in your data warehouse."

In our survey, the organizations having the most success with agents (the data leaders) provide their AI with access to more than 70% of enterprise data. That figure falls to 45% among the survey group as a whole, and even further to 30% or less among the data laggards – the group having less success with their agentic AI.

Overall, financial services providers struggle the most with data access, only providing AI with access to 33% of their data on average, followed by healthcare companies with an average of 40% (see Figure 6). In heavily regulated industries such as these, legal restrictions account for some data being out of reach.

Figure 6: On average, only 45% of enterprise data is accessible to AI, but this varies by industry



Source: MIT Technology Review Insights survey, 2026



Wanted: Integration, access, velocity, and context



Legacy data systems struggle to unlock dark data for a number of reasons. For example, data could be pocketed away in departmental or other silos, with no integration layer available to connect them with business context. Legacy infrastructure is also not designed for the scale and power agents require. And data may be in outdated formats or stored in obsolete devices.

Surveyed executives point to four factors that constrain their existing data platforms from supporting AI agents:

- Entrenched silos
- Difficulty accessing and managing unstructured data
- Insufficient access to real-time data
- A lack of business context and semantics

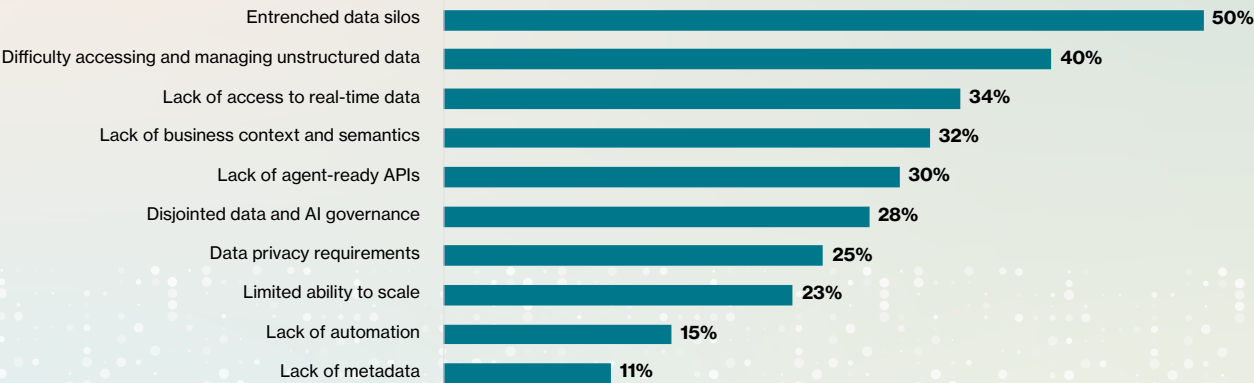
Fewer silos, greater access

The greatest challenge for surveyed organizations overall is entrenched data silos, as cited by 50% of all respondents (see Figure 7). This is most pronounced in manufacturing, retail, and media and entertainment, where nearly 60% of respondents cite this hurdle.

The second biggest challenge overall is accessing and managing unstructured data, as cited by 40% of all respondents.

For B2C organizations, a particular difficulty with unstructured data is analyzing customer information. “With structured data, it’s fairly easy to see personally identifiable information redactions,” says Mishra.

Figure 7: Unstructured data and silos are the biggest challenges for organizations with their existing data platforms



Source: MIT Technology Review Insights survey, 2026

"That's very difficult with unstructured data. Not many enterprises are experimenting with multimodal analytics yet because of this."

Real time and relevant

The third platform constraint, cited by 34%, is limited access to real-time data. To make decisions at speed, AI agents require frictionless, often instantaneous access to data. The batch-oriented architecture typical of older data platforms means that data is not updated continuously, making real-time access virtually impossible. "That's almost like you're giving your agent a view of the past and not how your business is doing today, because you don't have a current snapshot of data," says Bajwa.

A lack of business context and semantics is another significant challenge, cited by 32%. Context transforms data from being generally useful to being highly relevant to each individual use case. It is more than metadata, as vital as that is for bias mitigation and traceability purposes. When AI agents get the right context from a data system, they can understand not just where the data lives, but also what it means, how it is accessed, and the relationships between different data.

For AI agents to make good decisions, they need to understand how the data they consume maps to what is important for the teams or functional units that are using them and the goals they are pursuing. This level of context enables an agent to ignore data that is irrelevant to the user's goals and focus on relevant data.

"Context is a critical attribute of the data layer," says Thomas Hoy, director of clinical data management at HCA Healthcare. "A technically clean dataset is not enough. It also needs business definitions, ownership, lineage, quality signals, evidence, security, and usage constraints. When that context travels with a governed data product, users and AI-enabled applications can interpret the data more reliably."

Moving from data to knowledge at Deutsche Telekom

Making its data context-rich has been a priority focus of Deutsche Telekom's data platform team in recent years, according to Ashutosh Mishra, vice president of data strategy and platform at the company. That work has been two-pronged.

The first is lineage. "We've worked extensively on generating metadata, which gives us knowledge about physical entities or data assets," says Mishra. "Metadata has always been a problem area in the telecoms industry because of its unique taxonomy. So we've had to look hard at this." A first step was to build what Mishra calls "an agentic metadata generator," which essentially organizes the semantics in tables and columns.

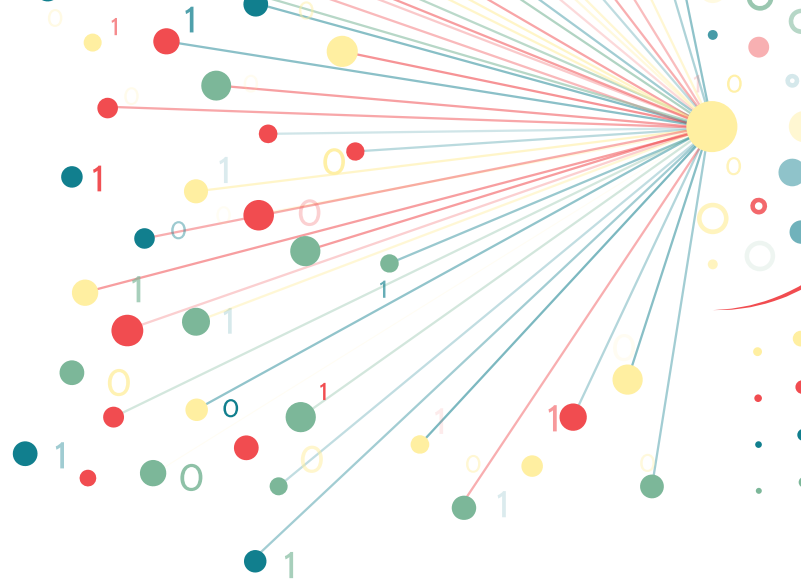
The second prong involves developing "an ontology of data," says Mishra. This entailed building a knowledge graph – a structured representation of facts about entities and their relationships – which, in turn, enabled Mishra's team to create a context graph. "This tells us how our data relates to business concepts. It can tell me how a customer is related to a product from a business concept point of view, and how the different types of products or product packaging are related to the customer."

Previously, says Mishra, users could not determine how relevant the data was to any particular conversation, whether 60%, 80%, or 90% relevant. "The knowledge and context graphs we've built have helped us to get much more value out of our data."

"Context is a critical attribute of the data layer."

Thomas Hoy, Director of Clinical Data Management, HCA Healthcare

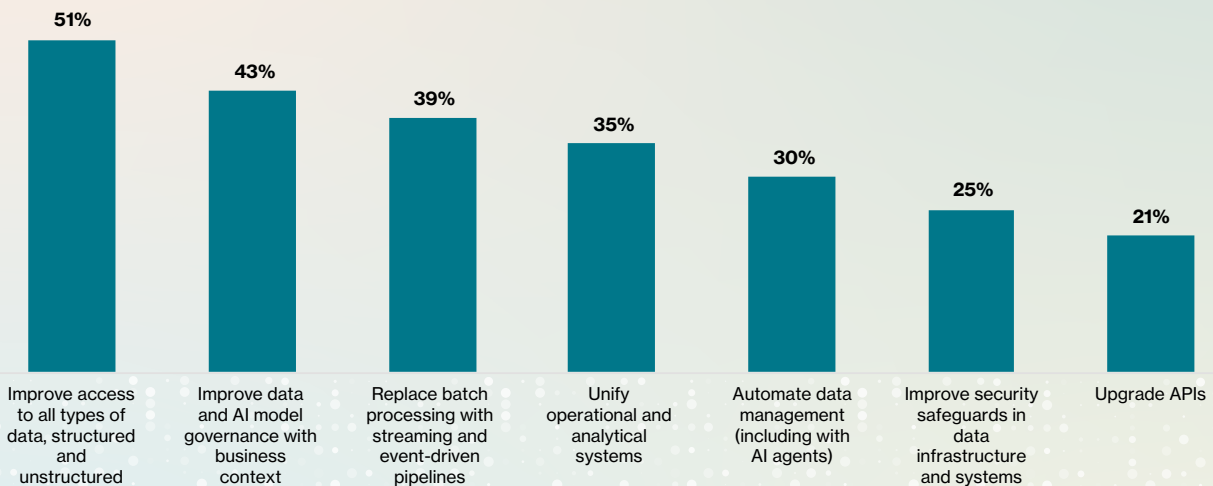
05 Building agentic readiness



How then do organizations address their legacy infrastructure constraints and evolve for the agentic era? Surveyed organizations are prioritizing two key initiatives. First, improving access to all data, both structured and unstructured. This is a priority for 51% of respondents overall (and 61% of retailers)

(see Figure 8). And second, enhancing data and AI governance with business context, as cited by 43% of respondents overall and a particular priority for those in healthcare and life sciences. Speed is another priority, with 39% intending to replace batch processing with streaming and event-driven pipelines.

Figure 8: Improving access to all types of data and enhancing business context are top priorities for scaling AI agents



Source: MIT Technology Review Insights survey, 2026

For data leaders, however, the priorities are somewhat different. At the top, cited by nearly twice as many leaders as in the overall group (56% vs. 30%) is the automation of data management with the use of AI agents (overall, high-tech organizations are much more likely to be pursuing this, too). This suggests that these companies are starting to move beyond agents executing individual tasks toward agents

managing entire workflows. Considerably more data leaders (40% vs. 21% of respondents overall) are also prioritizing the upgrading of their APIs, a necessary step to enable agents to orchestrate processes across different systems. The data leaders' focus on these initiatives indicates they are at a higher level of maturity in agentic AI use than most other organizations.

Building for scale

Priority data initiatives must be pursued with a view to enable agent scaling. This includes measures to improve context, which must be built to scale along with an organization's agents.

Scale also requires a large degree of agent autonomy. And for agents to be autonomous, they must be tightly connected with the organization's operational systems. Agents need to be able to perceive things happening in real time with live data so they can take action in the moment. More than a third of respondents (35%) say their organization plans to unify their operational and analytical systems.

An AI-native data platform will ensure all these attributes are in place to best support agents, says Gutmans. "A bolt-on architecture will be slow, expensive, brittle, hard to use, and, being a patchwork, will lead to spiraling costs, making it a financial liability," he says. "An AI-native data platform builds AI deep into the data processing and storage engines. That enables the most efficient data processing with AI."

HCA Healthcare's journey to an AI-native data estate

A little over a year ago, HCA Healthcare established a data office to build the trusted, well-governed foundation required for its AI strategy. According to chief data officer J.R. Allen, the office has three priorities: "First, we use a secure, scalable data lakehouse architecture on Google Cloud to integrate all enterprise data. Second, we make that data easier to use by applying governance so employees can understand the data and decide what is best for their needs. Third, we build reusable, well-governed data products to deliver this data in ways that support AI across HCA Healthcare."

One of the team's first uses of an AI agent was to generate definitions of all data. "We needed to generate these definitions at scale," Allen says. "We collect millions of data elements, and our data stewards could not have documented them quickly enough. We would have been constantly playing catch-up."

The agent accelerated that work by creating draft data definitions, which foundation models need to understand the meaning, context, and intended purpose of governed data assets. "We can now interact with those assets in natural language," Allen says. "We could not have achieved the scale we needed without an agent helping us generate the definitions. We used AI to make our data work for AI."

The data office has since expanded the use of AI-assisted methods across the organization. "AI can help strengthen the data ecosystem that downstream analytics and AI-enabled applications depend on," says Thomas Hoy, vice president of data architecture at HCA Healthcare. "For the data office, that means applying AI-assisted methods to activities such as metadata generation, documentation, classification, catalog publication, data-quality checks, and release readiness. These capabilities augment those of our data stewards and engineers and remain part of governed data management processes with accountable human review."

Hoy notes that standardized interfaces such as Model Context Protocol (MCP) can provide approved applications with a consistent way to access governed data services. "Our role is to provide consistent, secure, and governed access to trusted data," Hoy says. "The interface may be natural language, an API, an agent, or another data-to-model pattern, but trust has to come from the governed data foundation underneath that interaction."

Conclusion: A framework for success

When AI moves beyond answering questions to taking actions, it is more critical than ever to get it right. If AI agents fail to make accurate decisions and take relevant actions, the effects could be negative. If agents consistently get it right, they could transform company fortunes.

Our research indicates that enterprise adoption of agentic AI is on the cusp of an extraordinary acceleration. Within two years, all the surveyed organizations plan to be utilizing AI agents, the vast majority of them widely. This rapid transition from limited pilots to widespread, always-on operation makes resolving legacy data bottlenecks an immediate priority for businesses hoping to realize their expected ROI from agent deployment.

As organizations look to scale agentic AI across the enterprise, they cannot ignore their data systems. Ultimately, the ability of AI agents to make the right decisions – at speed – rests largely on two closely interlinked factors: data access and trusted context.

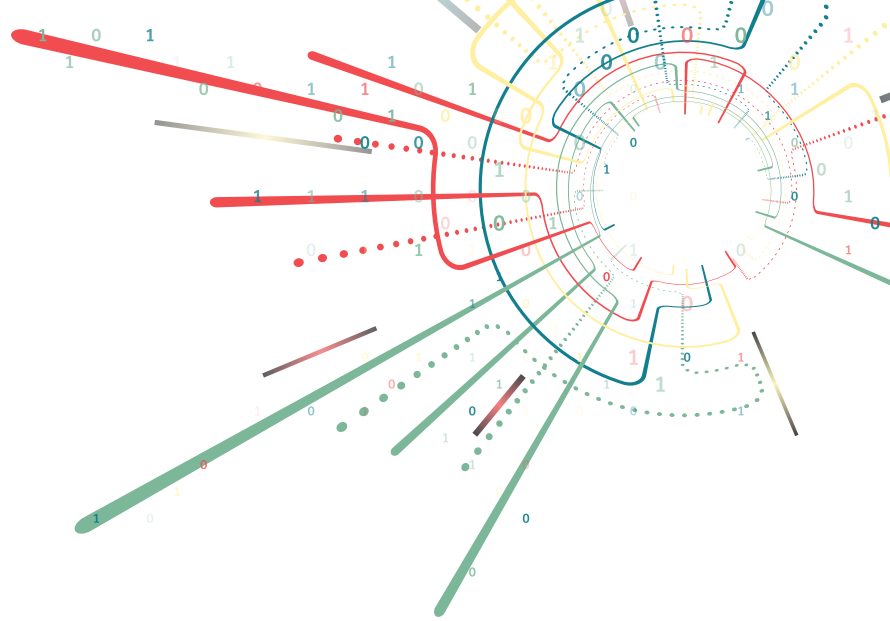
Getting the foundation right allows companies to develop trustworthy and reliable AI – and generate the promised returns.

“You have to be able to trust the agents’ outcomes,” says Gutmans. “If you can’t trust the outcomes, you’re still going to have a lot of human-in-the-loop manual validation for every automated action. That’s going to slow you down and impact your ROI.”

Our research has found that those having the most success with agentic AI – the data leaders – are building a firmer data foundation for agentic AI success than the majority. And legacy data systems are much less of a concern for them. By providing access to most of their data, these leaders are seeing successful outcomes with agents across many business functions. And they are finding it easier to scale and trust agents to make accurate decisions at speed.

“A bolt-on architecture will be slow, expensive, brittle, hard to use, and, being a patchwork, will lead to spiraling costs, making it a financial liability. An AI-native data platform builds AI deep into the data processing and storage engines. That enables the most efficient data processing with AI.”

Andi Gutmans, Vice President and General Manager, Data Cloud, Google Cloud



For other organizations looking to establish similarly AI-ready foundations, three lessons from the data leaders offer a place to start:

1. **Activate all of your data.** Organizations can start by discovering, inventorying, and classifying their unstructured and other dark data assets in addition to their structured data. After applying governance principles, information can be extracted and connected to AI agent use cases. Ultimately, achieving true agent autonomy at scale will require unifying operational and analytical systems.
2. **Put a premium on context.** Business context is what will allow organizations to activate their data in a highly trusted way. Solutions are now available that help organizations connect their fragmented data sources, linking them to AI agents and giving the data meaning. Organizations should be evaluating their options for delivering context now.
3. **Think AI-native.** AI-native systems are designed to operate in an AI environment and built from the outset to leverage AI for data management and decision-making. Initiatives to replace existing systems should focus on AI-native solutions.

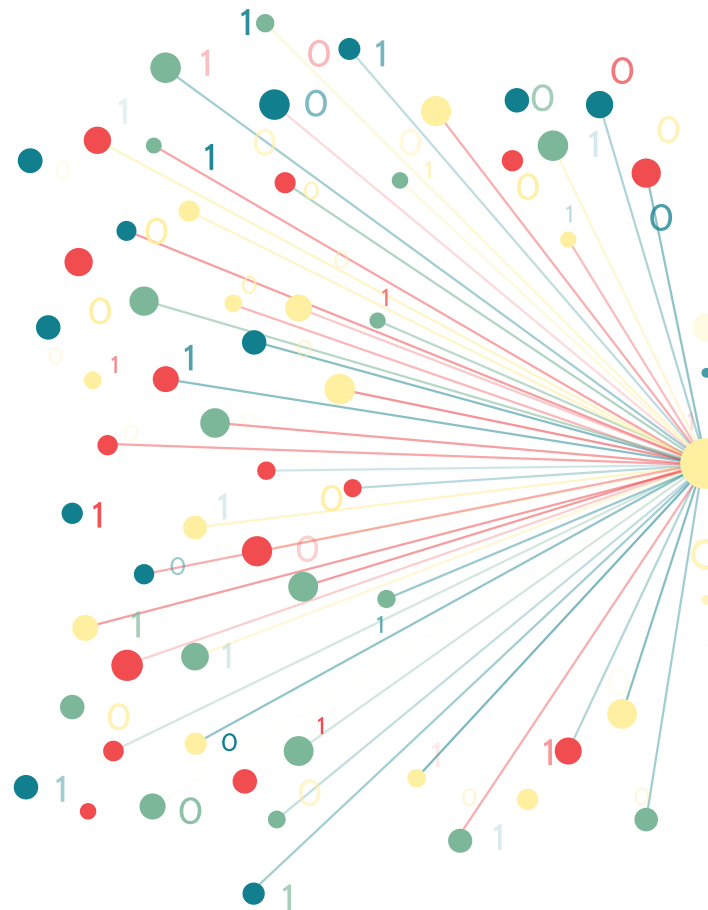
“When organizations want autonomous agents that not only reason on their behalf but also take autonomous action, the bar for trust and quality is much higher than before,” says Gutmans. “This is why it is so critical to choose the right data platform.”

“When organizations want autonomous agents that not only reason on their behalf but also take autonomous action, the bar for trust and quality is much higher than before. This is why it is so critical to choose the right data platform.”

Andi Gutmans, Vice President and General Manager, Data Cloud, Google Cloud

“Legacy hugely impacts ROI. Running AI agents with legacy data systems requires a lot of procurement, infrastructure cost, and engineers to maintain and develop them. That entails massive upfront investment.”

Ashutosh Mishra, Vice President, Data Strategy and Platform, Deutsche Telekom



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. This content was researched, designed, and written by human writers, editors, analysts, and illustrators. This includes the writing of surveys and collection of data for surveys. AI tools that may have been used were limited to secondary production processes that passed thorough human review.

About Google Cloud

Google Cloud offers a powerful, optimized AI stack – including AI, infrastructure, developer, data, security, and collaboration tools built for today and tomorrow. Google Cloud offers a powerful, fully integrated and optimized AI stack with its own planet-scale infrastructure, custom-built chips, generative AI models and development platform, as well as AI-powered applications, to help organizations transform. Customers in more than 200 countries and territories turn to Google Cloud as their trusted technology partner. To learn more about Agentic Data Cloud and how Google Cloud can help you build a trusted data foundation for the agentic era, please visit <https://cloud.google.com/data-cloud>.



Endnotes:

1. See, for example: "The state of AI in the enterprise", Deloitte, January 2026, <https://www.deloitte.com/uk/en/issues/generative-ai/state-of-ai-in-enterprise.html>; "The state of AI in 2025: Agents, innovation, and transformation", McKinsey & Company, November 5, 2025, <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>.
2. "AI ROI: The paradox of rising investment and elusive returns," Deloitte, October 22, 2025, <https://www.deloitte.com/uk/en/issues/generative-ai/ai-roi-the-paradox-of-rising-investment-and-elusive-returns.html>.
3. "Gartner Announces the Top Data & Analytics Predictions," Gartner, June 17, 2025, <https://www.gartner.com/en/newsroom/press-releases/2025-06-17-gartner-announces-top-data-and-analytics-predictions>.

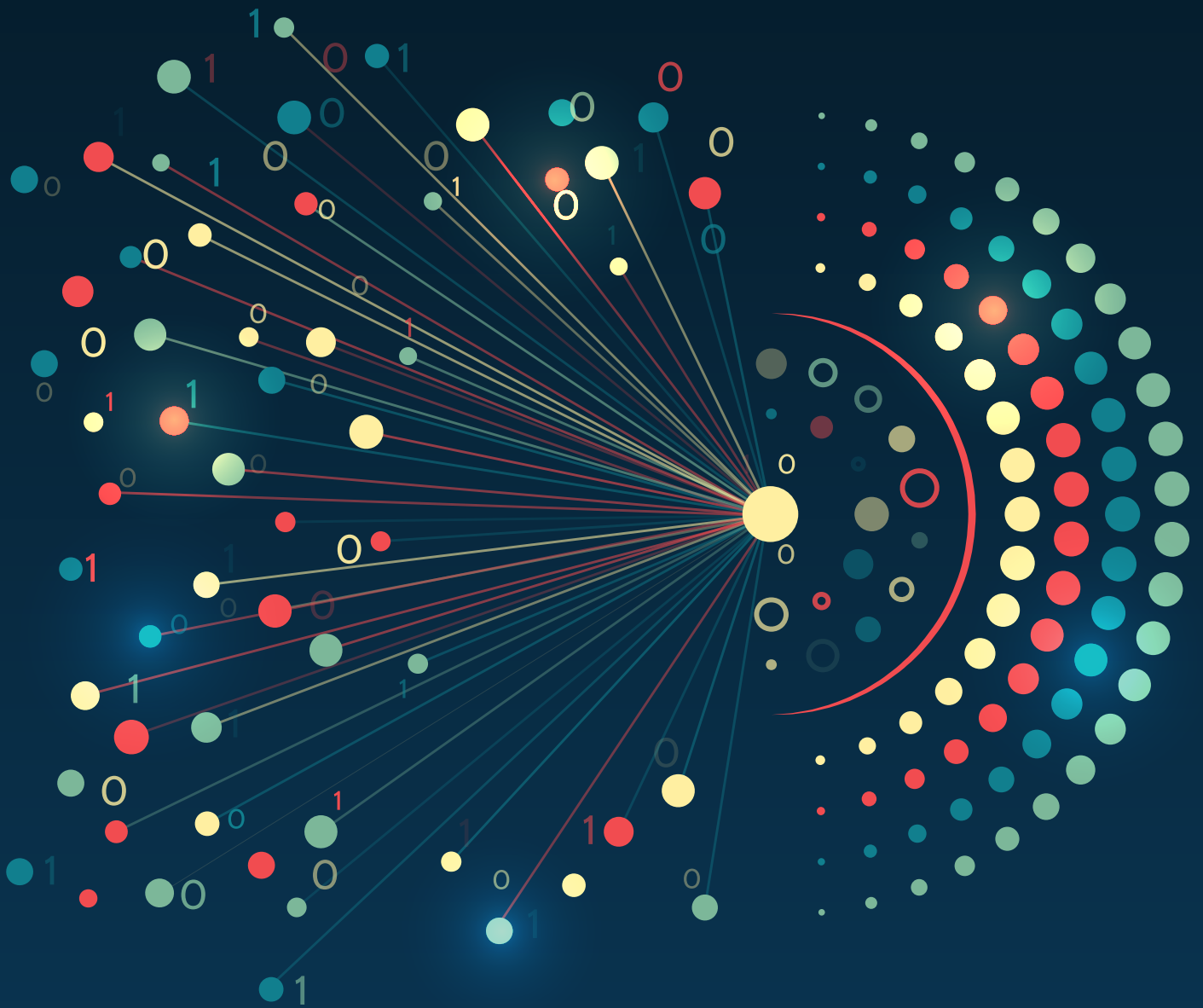
Illustrations

Illustrations assembled by Tim Huxford with elements from Shutterstock and Adobe Stock.

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.

© Copyright MIT Technology Review Insights, 2026. All rights reserved.

To cite this report, please use: "Scaling AI agents with trustworthy data," MIT Technology Review Insights and Google Cloud, August 2026.



MIT Technology Review Insights

www.technologyreview.com

insights@technologyreview.com