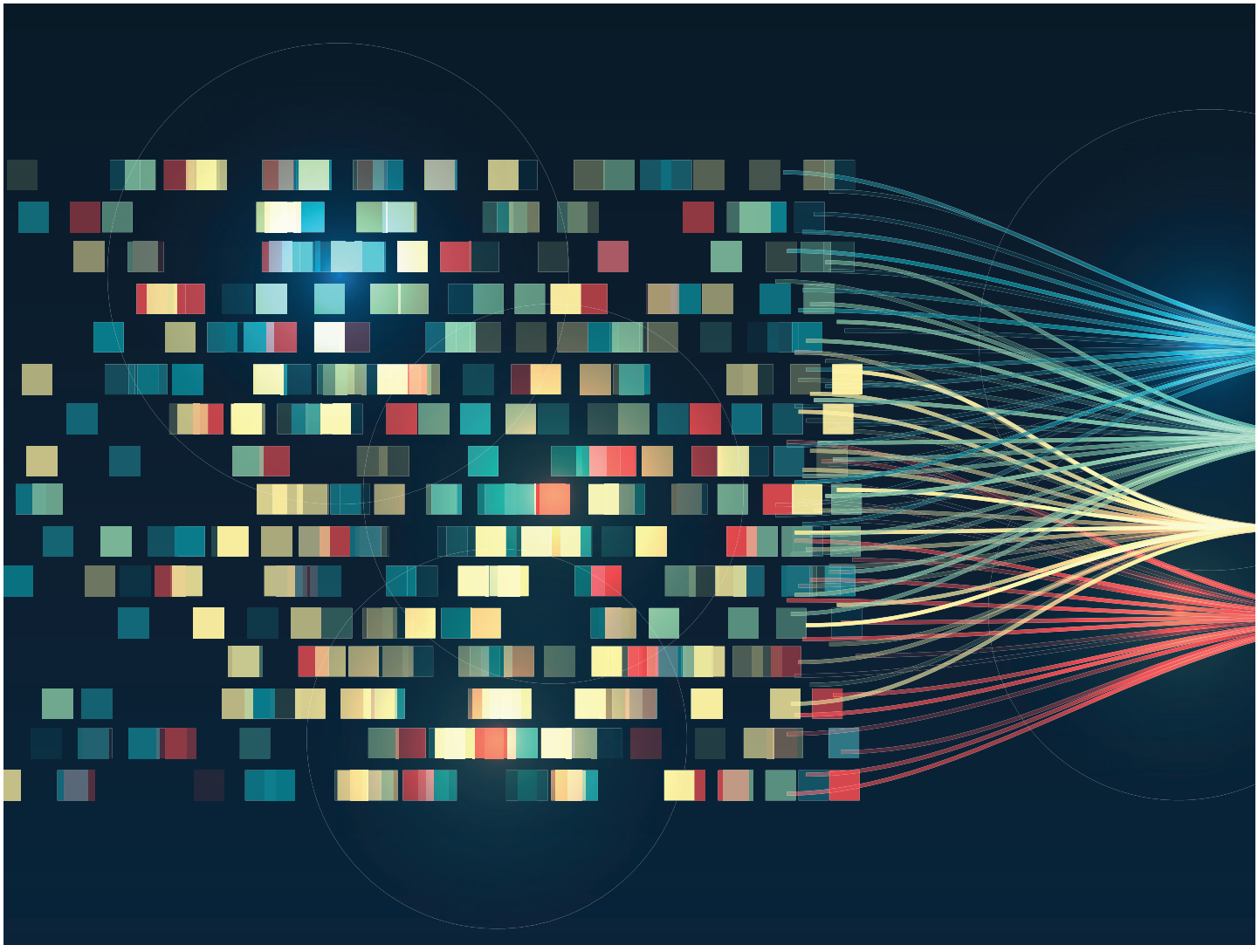
MIT Technology Review Insights

Produced in association with

kyndryl

More diverse data estates require a new strategy—and the infrastructure to support it.

Modern data architectures fuel innovation



Key takeaways

1. Inadequate data management has substantial costs, slowing companies' access to strategic insights and impeding the implementation of advanced technology such as AI.
2. A modernized data infrastructure can provide data users across the enterprise with one-stop, self-service access to the data they require to build business solutions.
3. An emerging type of data architecture, called a data fabric, integrates and orchestrates data from numerous sources. Built in composable layers, a data fabric allows an organization to prioritize the data functions of most strategic importance.

Companies have contended with a deluge of data for years. And while most have not yet found a good way of managing it all, the challenges – diverse data sources, types, and structures and new environments and platforms – have grown ever more complex. At the same time, deriving value from data has become a business imperative, making the consequences of not managing your organization's data more severe – from lack of critical business insights to the hobbling of AI implementations.

Greater data complexity leads to greater consequences

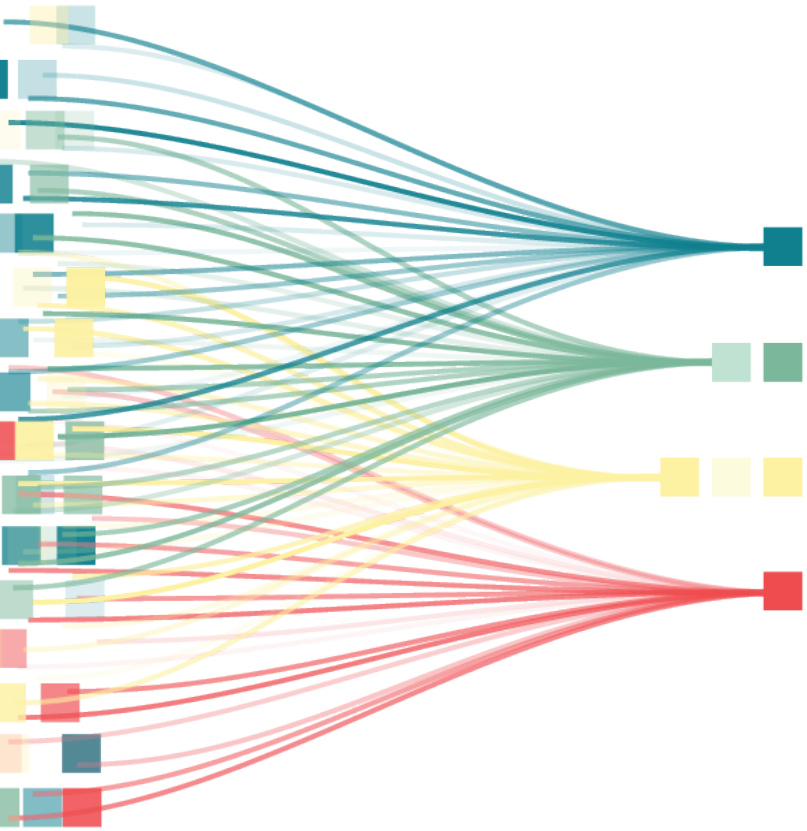
Data is not only increasing in volume, velocity, and variety, but also the data estate has become increasingly intricate. For years, organizations have struggled with data being sequestered in separate silos within the company. Today, data location adds another layer of complexity, with some of the data on premises, some of it in the cloud, and some of it coming in streams from the edge. By 2025, **more than 50%** of enterprise-critical data will be created and processed outside the data center or cloud, Gartner analysts estimate. In order to be truly data-driven, organizations realize, they must reach both wider and deeper into their operations, identifying and digesting data and information from various departments and sources.

"Each line of business is driving digital transformation in its own way," says Naveen Kamat, executive director and CTO of **data and AI services** at Kyndryl, an IT infrastructure services provider. "They are setting up their own apps in the cloud, which generate data daily. Then there's web and social media data coming in. The enterprise data estate is becoming much, much bigger; it's becoming much more complex to manage."

The insurance industry provides an example of today's data landscape complexity. One substantial challenge to good data management in insurance is a plethora of legacy systems built up over the years, says Ali

"The enterprise data estate is becoming much, much bigger; it's becoming much more complex to manage."

Naveen Kamat, CTO of data and AI services, Kyndryl



Shahkarami, chief data officer at Allianz Global Corporate & Specialty (AGCS). "That's especially true for international companies operating across borders with different products, regulatory requirements, and reporting requirements," he notes. "The ability to do that centrally and in a consistent manner is a big challenge. It impacts everything you build with data and analytics."

Unfortunately, while data management has become more challenging, data management skills have become harder to come by. The number of skilled data personnel **has stayed the same or even dropped** over the last decade, even as the number of data and application silos have increased, says Gartner. That means it takes more time than ever to meet integrated data analytics needs.

The consequences for organizations that fail to manage their data effectively and efficiently are becoming dire. For one thing, the cost of inadequate data management is growing. The cost of poor data can be about 20% of revenue, estimated Thomas C. Redman, president of consultancy Data Quality Solutions, in a co-authored **MIT Sloan Management Review article**.

"Almost all work is plagued by bad data," write Redman and Thomas H. Davenport. "The salesperson who corrects errors in data received from marketing, the data scientist who spends 80% of his or her time wrangling

The path to a modern data platform

When an organization is ready to build a modern data platform, how should it get started? Naveen Kamat, executive director and CTO of data and AI services at Kyndryl, advises the following steps:

Find your data. The first step is to understand what data the organization has and where it is located. On average, 55% of an organization's data is "dark" – unknown, unused, or unfindable – according to research by TRUE Global Intelligence. "Because it's coming in from so many sources, enterprises struggle to get a view of all of their data," says Kamat.

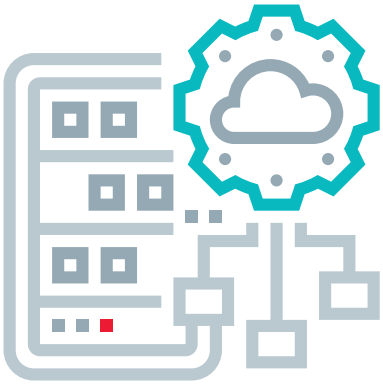
Determine your business goals and data strategy. "We look at their data strategy," Kamat says. "What are their business goals? Do they have the data to meet those goals?" That in turn dictates the choice of data architecture that will support discovery and make data easy to use.

Harmonize and integrate data. Today's modern architectures accommodate many sources and different types of data. Data doesn't necessarily need to be collected in one place. Rather, a modern federated architecture enables creation of pathways from silos, the cloud, and the edge to allow data to be distributed easily, so users can tap into data wherever it lives.

Doing so, however, requires a common vocabulary. "You have to agree on a shared language," says Thomas C. Redman, president of consultancy Data Quality Solutions. Kamat agrees: "You need a common taxonomy that works across lines of business and functional teams." If what you pull from different sources doesn't refer to the same data, a machine learning model could be ineffective at best, or a disaster at worst.

Apply governance. Although the goal is to make data accessible, you also need guardrails. "You want to democratize data access, but you also need to make sure access is secure, that only those with the right level of authorization can access certain data," says Kamat.

Create data products. Once you have created pathways and established a common taxonomy, you can build data products for use across the organization.



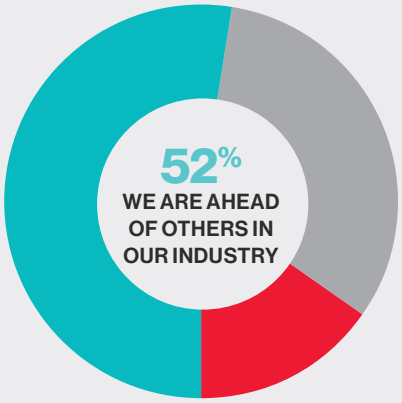
Redman and Davenport estimate that less than 5% of companies use their data and data science to gain a competitive edge.

data, the finance team that spends three-quarters of its time reconciling reports, the decision maker who doesn't believe the numbers and instructs his or her staff to validate them."

Redman and Davenport estimate that less than 5% of companies use their data and data science to gain a competitive edge. "Companies are not seizing the strategic potential in their data," they conclude.

When it comes to implementing advanced technologies, such as machine learning and artificial intelligence, inadequate data management represents a substantial barrier. Not only could AI programs be ineffective, but "without the right data, building AI is risky and possibly dangerous" if data bias, diversity, and systematic labeling are not part of a data management strategy, **says Rita Sallam**, distinguished vice president and analyst at Gartner.

C-level executives are moderately confident in their data strategy execution



Compiled by MIT Technology Review Insights based on data from Frost & Sullivan 2022 Data Management and AI Survey

Modern data management requires a data-driven strategy

At its most basic level, the problem is this: enterprises have outgrown their data management strategies and architectures. To handle today's challenges, they need to modernize.

Building a modern, effective data strategy requires two things: strong executive leadership and a strategic choice of technology. First, leadership must make a true commitment to creating a data-driven culture.

"Management still doesn't realize that this is not primarily a technology issue," Redman says. "To solve the problem, you've got to properly diagnose it. And most organizations [still] diagnose this as a technical problem. It is not a technical problem." Kamat adds that he often sees companies trying to tackle the technical problems first, to their detriment. "Organizations often struggle to integrate data from legacy systems, but sometimes it can be legacy thinking that gets in the way," he says. "You need leadership within the organization to look at and have the determination to be a data-driven enterprise. "

Once they have committed leadership, organizations can formulate a strategy and adopt technology to modernize their data infrastructure. A modern architecture should enable companies to locate their data, harmonize and integrate data from many sources, apply the appropriate governance, create data products that can be used across the organization, and automate data engineering tasks to reduce complexity, minimize costs, and optimize business value.

The Allianz corporate insurance carrier blends strong leadership and the right technology through a global data office that works with both IT and lines of business. "That three-way collaboration – between the business, IT, and [the global data office] – is key," says Shahkarami. "This allows us to deploy solutions that bring business value."

The importance of a shared data language

Before it can have a true data-driven culture, an organization requires a common language. But convincing senior leadership of that can be challenging, says Thomas C. Redman, president of consultancy Data Quality Solutions.

Establishing a lingua franca doesn't mean various departments can't continue to speak in their native tongues, notes Redman. "Within your specialty, it's perfectly fine to have your own language and your own way of doing things—it really builds efficiency," he explains. "But when you work across departments, you need to agree on a shared language." And that language needs only a limited vocabulary to be effective. "If you do this at the right level of abstraction, you only have to agree on maybe a hundred key terms," he says.

To demonstrate the need for common definitions, Redman takes senior leaders through a simple exercise. "Hand everybody in the room a piece of paper and ask them to write down their definition of a few words, like 'customer,' 'profit,' and 'product,'" he says. Then collect the papers and read off the definitions. Participants may be astounded to learn that "customer" means a qualified lead to sales, while in finance it refers to whoever is responsible for paying the bill. The definitions can be so different that they sound like they came from different companies.

While the exercise is good fun, Redman cautions that these differing definitions, if not harmonized before launching a data project, will proliferate through the new architecture and cause problems in the future. "Three years down the road, when you're in the midst of a merger or you're implementing advanced AI, you're in trouble," he says.

Solving business problems with a 'one-stop data shop'

Those solutions impact the bottom line. With more and more information and operations becoming digitized, analyzing and acting on the data produced is the key to competitiveness. In manufacturing, for example, factories now generate reams of valuable internet of things (IoT) data, from sensors, systems that monitor and control industrial processes, and other sources. With a modern data architecture, this data is available to be used to predict potential problems – such as a piece of equipment that's about to fail – and proactively address them, thus reducing unplanned outages. By fixing things before they break, manufacturers can improve operational efficiency.

AGCS is modernizing its architecture in ways that benefit both internal and external insurance customers, says Shahkarami. He uses an analogy: just as drivers should be able to fill up at a gas station rather than refine their own oil, his customers should have simple on-demand access to the data they need in the form they require. Internally, a modern architecture simplifies data access and analytics so internal users can make underwriting, claims, pricing, and actuarial decisions on a self-service basis. "We are creating a seamless one-stop shop of data so anyone can create a solution," he says. Externally, the company is using its data to improve customer service: speeding up the delivery of quotes to brokers, helping salespeople answer questions more quickly, making the claims adjustment process faster and more efficient. Says Shahkarami, "We are creating a pipeline of use cases that solve real business problems."

Weaving an enterprise-wide data fabric

The most effective means of building a modern data infrastructure is an emerging type of architecture, called a data fabric, that integrates and orchestrates data from many sources across a distributed environment. Typically agnostic to the various data processes, structures, and locations it integrates, the

fabric acts as connective tissue among them. It uses metadata to make data available throughout the organization in an automated way.

A data fabric enables users to access and consume data from across the organization in a self-service fashion. It also enables the organization to automate data governance, data privacy and security, data engineering work, and data integration.

According to Gartner, a data fabric can reduce data management tasks, including design, deployment and operations, by up to 70%. **The firm predicts** that, by 2024, data fabric deployments will quadruple efficiency in data utilization, while cutting human-driven data management tasks in half.

Data fabrics can be built in distinct, composable layers – each consisting of a function such as data collection, ingestion, integration, or governance – that give organizations the flexibility to prioritize implementing layers that are key to their strategy or return on investment. "It's like a set of Lego blocks, so the customer can focus on whatever they need to at their stage of digital transformation," says Kamat.

Enabling emerging technologies

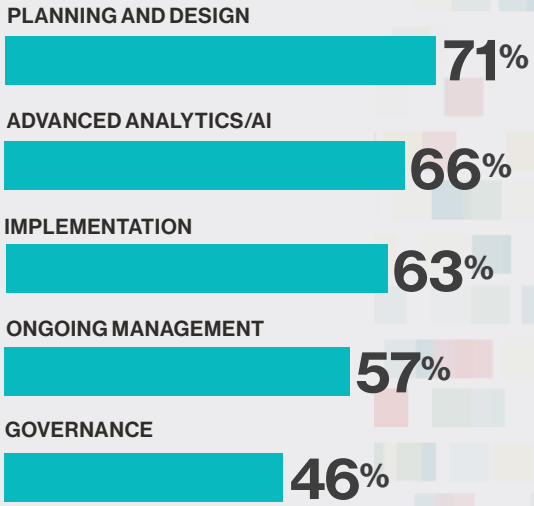
Perhaps most important reason to modernize is to enable effective machine learning and AI. Once a composable data fabric is in place, organizations can not only automate much of their data management, but they also can implement and automatically manage AI using commercially available machine learning operations (MLOps) solutions. MLOps products can, for example, monitor models and alert you when they need fresh data and retraining. Financial operations models are available to help manage spending on cloud.

As modern architectures pave the way for better management and analysis of data, AI will help businesses see around the next corner, helping them better prepare for the future. Shahkarami returns to an automotive analogy to explain AGCS's plans: "We are trying to look ahead rather than only looking in the rearview mirror," he says. "Today's cars can sense an accident about to happen and warn the driver." In Allianz's corporate business, he continues, "using weather forecasts and other data, our modeling capabilities can predict the impacts from a hurricane during or prior to landfall, helping Allianz better manage its business and serve its customers. In addition, risks such as cyber exposures evolve so quickly that an analysis of historic claims data would only offer limited insights into today's or even tomorrow's risks."

"There are enterprises that we see that are completely data-driven and really use AI in their operations," says Kamat. "They're now looking at newer things to bring in and newer emerging technologies, like blockchain and metaverse. They're on the cutting edge, and they're exploring the latest and the best in this space." Organizations that modernize their data management architectures will be well positioned not only to navigate their current competitive landscape, but also to continually adopt AI and other emerging technologies to maintain that competitive edge.

Organizations are calling on outside help

% using a third-party provider for each task



Compiled by MIT Technology Review Insights based on data from Frost & Sullivan 2022 Data Management and AI Survey

"We are creating a seamless one-stop shop of data so anyone can create a solution."

Ali Shahkarami, chief data officer, Allianz Global Corporate & Specialty

"Modern data architectures fuel innovation" is an executive briefing paper by MIT Technology Review Insights. We would like to thank all participants as well as the sponsor, Kyndryl. MIT Technology Review Insights has collected and reported on all findings contained in this paper independently, regardless of participation or sponsorship. Laurel Ruma and Teresa Elsey were the editors of this report, and Nicola Crepaldi was the publisher.

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 U.S. 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

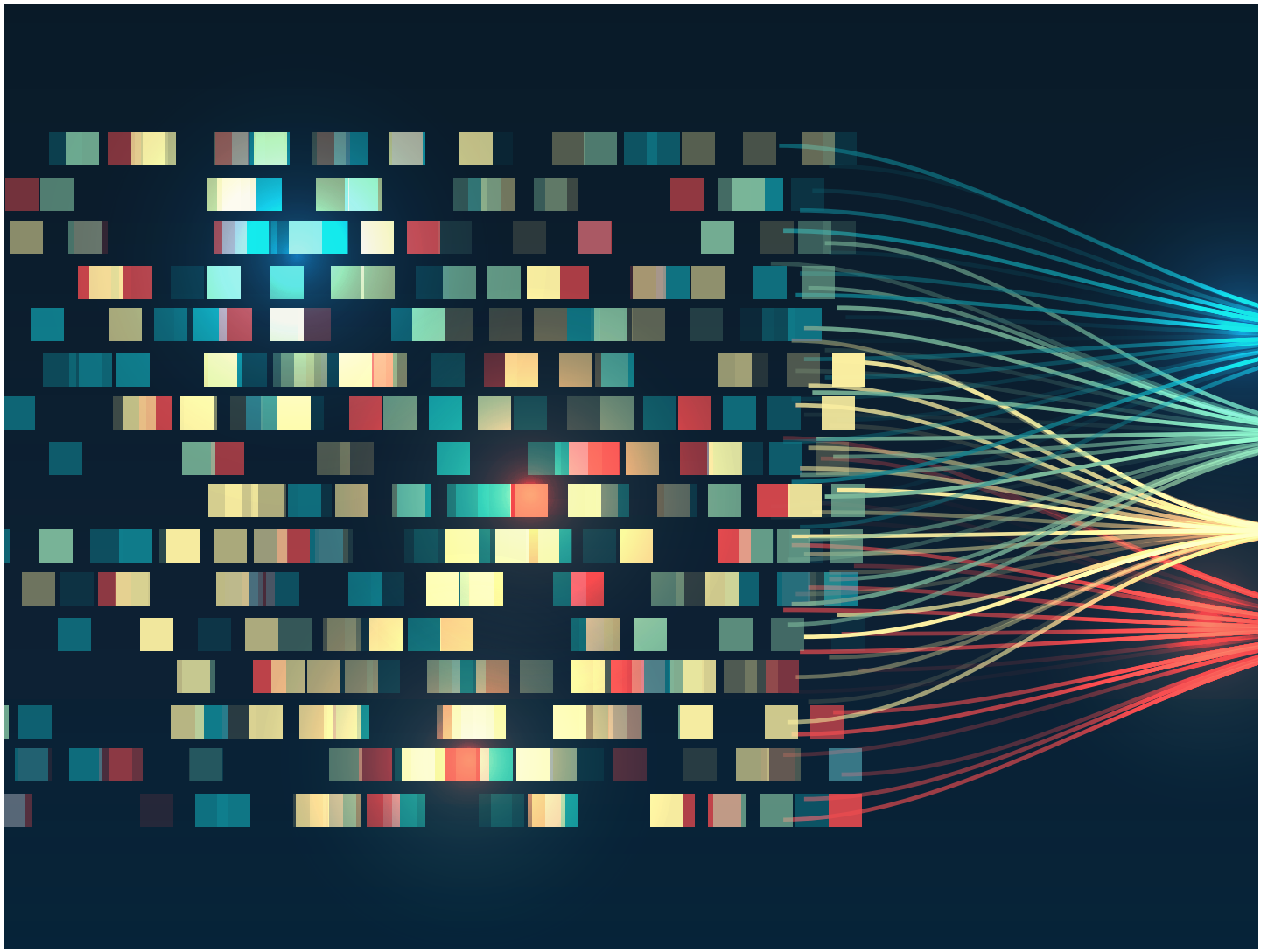
Kyndryl (NYSE: KD) is the world's largest IT infrastructure services provider, serving thousands of enterprise customers in more than 60 countries. The company designs, builds, manages and modernizes the complex, mission-critical information systems that the world depends on every day. For more information, visit **www.kyndryl.com**.

kyndryl™

Illustrations
Cover art by Adobe Stock and spot illustrations created by Chandra Tallman with icons by The Noun Project 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 on any person in this report or any of the information, opinions, or conclusions set out in this report.

© Copyright MIT Technology Review Insights, 2023. All rights reserved.



MIT Technology Review Insights

www.technologyreview.com

@techreview @mit_insights

insights@technologyreview.com