



Data readiness in the age of generative AI



**All business leaders need a greater
understanding of the importance of
data in the age of generative AI.**

Six data essentials to stay ahead of the pack

Generative AI stands poised to help companies reinvent themselves—by streamlining operations, delivering better customer experiences and otherwise driving growth in many ways. That's why, in a global survey by Accenture of 2,300 business leaders **nearly all (97%) said they believe that generative AI is a “game-changing” technology worthy of long-term investment.¹**

However, despite these great expectations, many companies struggle to turn their generative AI pilots into scalable solutions. A big reason for this, our research also shows, is that most organizations are still not “data ready” for generative AI. For example, when Accenture surveyed 2,000 CXOs in 2024, **48% said their organizations lacked enough high-quality data to operationalize their generative AI initiatives.²**

So what does it mean to be data-ready for generative AI? Companies with high readiness have the right data, with the right quality, in the right quantity. They meticulously capture data from every aspect of their business and curate it so they can analyze and use it later. And they have robust governance to manage, maintain and operate their data responsibly.

To their credit, most companies today recognize the value of data and have established multiyear programs to enhance their data capabilities. Indeed, in the same 2024 survey, **75% of executives said that “good quality data” is the most valuable ingredient to enhance their generative AI capabilities—now and over the next six months.³**

In Accenture's 2023 report, [Cloud data: A new dawn for dormant data](#) we discussed the steps for unlocking data value. Those steps still hold true. But given the seismic shift in the data and AI landscape since the emergence of generative AI, the time is right to explore how things have changed.

This report—which draws on insights from approximately 1,800+ generative AI client projects, as well as Accenture's latest research—highlights 6 essentials about data and its role in powering reinvention:

- 1. Proprietary data is a competitive advantage**
- 2. Unstructured data holds untapped potential**
- 3. Synthetic data fills gaps in real-world data**
- 4. Connected data is key to context for generative AI**
- 5. Generative AI accelerates data risks**
- 6. Generative AI jumpstarts data readiness**

To help companies achieve a foundational level of data-readiness, we conclude with key actions that all data-focused executives should be taking.

1

Proprietary data is a competitive advantage

Generative AI tends to be most useful when it's powered by a company's proprietary data. That's because foundation models that run on company data can better unlock high-value insights about a firm's customers, products and operations; historical and real-time institutional knowledge can improve internal decision-making, reduce risks and identify new efficiencies, as well as open attractive monetization opportunities.

Though proprietary data is incredibly valuable, companies struggle to capture and harness it. That's why investing vigorously in proprietary data is essential, whether launching generative AI pilots or trying to scale them. To determine the correct level of investment, as well as to develop and maintain accurate data, firms should treat their data as a product (what we call a "value-led" approach).⁴

Treating data as a product requires identifying the unique data generated at each step of an organization's business processes and then selecting the data that's needed to

differentiate decision-making for different purposes. In certain situations, proprietary data (first party data) can also be complemented by data supplied by partners (second party) and by data collected externally (third party). A value-led approach to data allows companies to find the right balance between proprietary and non-proprietary data when deploying generative AI.

As part of its ongoing reinvention journey, global financial services group BBVA partnered with Accenture to develop a new, comprehensive digital sales model. The result? Nearly 50 million customers now interact with the bank through digital channels, and seven out of 10 sales are made digitally. BBVA's client onboarding process takes just minutes (versus a few days at most other banks), using AI-based

facial recognition and text analytics to verify account applicants via mobile app and real-time connections to external data sources to detect fraud.

By combining first-party data with new data sources to deliver a step-by-step view of the customer journey, BBVA's new digital sales model helps the bank prioritize sales initiatives for new customers and cross-sell to existing customers. The new model incorporates strategy and planning, paid media, search engine optimization, marketing automation, analytics and content production for BBVA's digital channels to reach individuals in hyper-personalized ways.

2

Unstructured data holds untapped potential

Companies' data tends to be structured (having a predefined format, such as being ordered in rows and tables), which creates a "pre-conceived view" of information. In contrast, unstructured data (encompassing formats like text, images, audio and video) is rich with contextual information.

That's why unstructured data has so much potential: It provides a real-life, unfiltered representation of a company's business. It's also why nearly three out of four companies that we surveyed are leveraging complex, real-time, unstructured data from multiple sources.⁵

When combined with structured data, unstructured data adds the context needed for more human-like communication from generative AI. Unstructured data, for instance, contains signals for tone, personality, look and feel that—when infused into foundation models—drives much richer interactions between people and machines.

For many years, Fortune has rigorously collected and analyzed complex financial data on the largest companies in both the US and around the world, in order to create the iconic Fortune 500® and Fortune Global 500™ lists. Together, Accenture and Fortune collaborated to transform that business knowledge into a Fortune Analytics™ LLM tool—an intuitive, user-friendly, generative AI-powered platform that provides access to insights from the Fortune 500® rankings. The platform is powered by Accenture's foundation-model services and proprietary large language model (LLM) assets and is fine-tuned with comprehensive Fortune datasets.

Users can receive useful graphical data visualizations like scatterplots, line charts and bar charts—generated on demand by the large language model based on the user request.

To unlock the potential of unstructured data, however, companies need to increase its accessibility and availability. Doing this, in turn, requires things like extending data architectures, enhancing data security and strengthening data governance. For example, companies could deploy scalable systems to manage unstructured data (such as data lakes to store and vector databases to serve data), develop real-time streaming analytics capabilities, integrate AI-driven tools for data classification and search and enforce strict-access controls to protect sensitive information.

3

Synthetic data fills gaps in real-world data

Generative AI is hungry for data: the more complex the task or output, the more data is required, in both quantity and quality. Synthetic data—created by algorithms, rather than collected from actual life events—addresses the growing scarcity of specialized datasets, enabling companies to explore multiple scenarios without the extensive costs associated with collecting traditional data.

For example, a company might use synthetic data derived from its products and customers to save time and money during market testing. Synthetic data can also be used for risk-management, for designing “what-if” scenarios and for removing bias from data.

Synthetic data can address certain data risks as well. If an organization’s real-world data is confidential, such as with medical records, the synthetic kind can be used to train AI models while protecting patients’ privacy. When regulations require data to be stored for long periods, keeping copies of synthetic data (rather than the original)

reduces the damage in the event of a cyberattack or other data breach. Despite these benefits, only half of the companies we surveyed have access to and can leverage synthetic data for their models.⁶

To make the most of synthetic data, companies need domain expertise to create and use the technology to strike a balance between quantity and quality. Likewise, companies require access to sophisticated frameworks that enable them to confirm that the synthetic data they created turned out as intended.⁷

Digital twins and expert knowledge can also fill in the data gaps needed to create an AI-powered supply chain. Say, for example, in an automated warehouse an automated guided vehicle (AGV) gets stuck, products are unavailable or equipment is missing. A lack of data means companies can’t see what is happening and what could be done about it.

Digital twins can simulate warehouse operations that are validated with operational data and employee know-how to power AI that predicts the next best action when operation anomalies occur. These twins also serve to enable what-if analysis for simulating new scenarios and to validate AI recommendations to see the impacts in throughput in the warehouse and reduction of waste. This paves the way for generative AI “agents” that can research, plan and recommend a course of action.

4

Connected data is key to context for generative AI

Turning enterprise data into insights entails sharing deep subject matter expertise between many people across an organization. Yet companies struggle to contextualize and find new relationships in data because much of it is locked in silos and functional domains. Indeed, **65% of CXOs in the Accenture survey referenced above said that building an end-to-end data foundation was one of the top obstacles to scaling generative AI. This end-to-end data foundation breaks down silos and makes quality data available by managing the entire data lifecycle—from initial collection to post-use management.⁸**

The unhappy result is that turning data into insights often takes days, weeks or months. Fortunately, generative AI can dramatically shorten that time frame to minutes or less.

That's what Accenture is doing with BMW, using our generative AI platform EKHO (Enterprise Knowledge Harmonizer and Orchestrator) to collect and analyze its enterprise data. The platform uses large language models to intelligently answer complex questions across business functions and use cases. The heart of the platform contains multiple AI-enabled applications (GPT agents) that intelligently choose the right data source and pull information based on the user's questions and enterprise-specific data. Thanks to the platform's flexibility, EKHO can be applied to a vast number of tasks across the company—and on the showroom floor.

In these and other ways, businesses can apply generative AI to break down data silos and discover more efficient

ways of working. To achieve this, every part of the organization must make data accessible and treat it as a valuable product—reliable, secure and easy to use.

Companies should also invest in the architecture and operating model needed to create, use and manage their data products. For example, creating a “semantic” layer can help a firm organize and define its data against business concepts, in a way that makes it easier for both people and generative AI to understand and engage with the data.

A retail firm, say, might have a huge dataset of customer interactions, sales records and product details. Though it would store its raw data in databases, the data would not be cross-functional and would be very difficult to interpret. By creating a semantic layer, sales reps could easily search for data using terms like “total sales in Q1.” And because the data would be presented with context (such as how customer-satisfaction scores relate to sales volumes), generative AI could analyze trends and make accurate predictions, too.

5

Generative AI accelerates data risks

Generative AI offers tremendous promise for companies, but it also creates and accelerates data-related risks. These can be legal, reputational or both. And they can touch on areas such as quality, privacy, security, bias, discrimination and intellectual property, among others. In December 2023, for instance, the New York Times sued OpenAI for training its models on the newspaper's articles without permission.⁹

Data-related risks can emerge from many directions. Generative AI makes data and AI tools more accessible, but often lacks safeguards against human error. Generative AI can also be intentionally misused to cause harm, such as by creating deep fakes, "poisoning" data and de-anonymizing data.

No wonder **42% of organizations in the same Accenture survey said they need help developing policies, governance and risk management processes for the responsible use of generative AI systems, ensuring compliance with regulations and laws.¹⁰**

And because laws and regulations on data and generative AI vary by jurisdiction and are evolving fast, legal compliance may become even more challenging in the years ahead. For example, the EU Artificial Intelligence Act, which enters into full effect in 2026, places AI systems into one of three categories and then regulates them accordingly (a total ban, extensive regulation or moderate regulation).¹¹

To mitigate these risks, companies should adopt robust data governance—something that is often baked into Responsible AI programs. Accenture's own internal Responsible AI program, for instance, has four main components.

The first involves raising leadership awareness about data and other AI-related risks, establishing Responsible AI principles and policies and setting up a dedicated Responsible AI team.

The second component involves conducting preliminary risk assessments and regulatory/enforcement reviews. The third component involves implementing standards for developing and purchasing AI, embed rigorous controls into the firm's technology, processes and systems and develop testing tools and persona-based training for employees. The fourth component involves ongoing monitoring and compliance of AI applications throughout their lifecycle.

Accenture is sharing best practices from its Responsible AI program with other organizations, too. The company recently partnered with S&P Global to train the financial-data provider's 35,000 employees on how to scale and innovate with generative AI, while using the technology responsibly.¹²

6

Generative AI jumpstarts data readiness

It's not just about what data can do for generative AI, it's also about what generative AI can do for data. Applying the technology to a firm's data processes can enhance various aspects of its data supply chain, from capture and curation to consumption. Generative AI can summarize and classify business data requirements, design documents and test cases, and generate runbooks and deployment scripts. Generative AI can, as noted, create synthetic data as well.

There are many other opportunities to apply generative AI to data migration and modernization programs. For example, teams of AI agents might automate the tasks of rewriting and improving software code. One agent could coordinate workflow. Another might handle code conversion. Yet another could focus on explaining how both the original and new systems work. (As with teams of human workers, teams of AI agents can often complete tasks more efficiently and effectively together than in isolation.)

Applying generative AI broadly across a company's data supply chain requires maintaining an ongoing knowledge base of things like metadata ("data about the data"), descriptive labels for different datasets and service tickets that track changes made to data over time. When the data lifecycle (how data is created, processed, stored and used) is eventually transformed by generative AI, companies need to update their data governance practices to ensure that their data remains trustworthy and otherwise well-managed. This includes creating new rules for handling AI-generated data ensuring that data remains consistent and verifying that it meets high-quality standards.

The road to data readiness

As we can see from these new data essentials, there is work for every organization to do when it comes to improving their data readiness in the age of generative AI. Generative AI has changed the road to data readiness whether you have been investing in data all along or are just getting started.

“Agentic generative AI workflows—coupled with responsible AI frameworks and centralized data management and governance—are the path to success, enabling seamless interaction with enterprise data to deliver insights and next best actions via a click or voice command. PepsiCo is well on the path to realize this vision to build the next gold standard in data management in the age of generative AI.”

Magesh Bhagavati

SVP and Global Head of Data, Analytics and AI, PepsiCo



It can be confusing to know where to start. We recommend the following actions to build a foundation of data readiness and prepare your data capabilities for generative AI.

1

Understand the value of your data and invest to maximize its full potential

In the era of generative AI, you already own your most valuable asset: your data. Relevancy with foundation models comes from your data.

Start by identifying the unique data that can be generated at each step of your business processes and the data that's needed to differentiate decision making.

Make sure to understand the economics of data so that you can better prioritize your investments required to capture and use your data. For example, storage, consumption and compute costs.

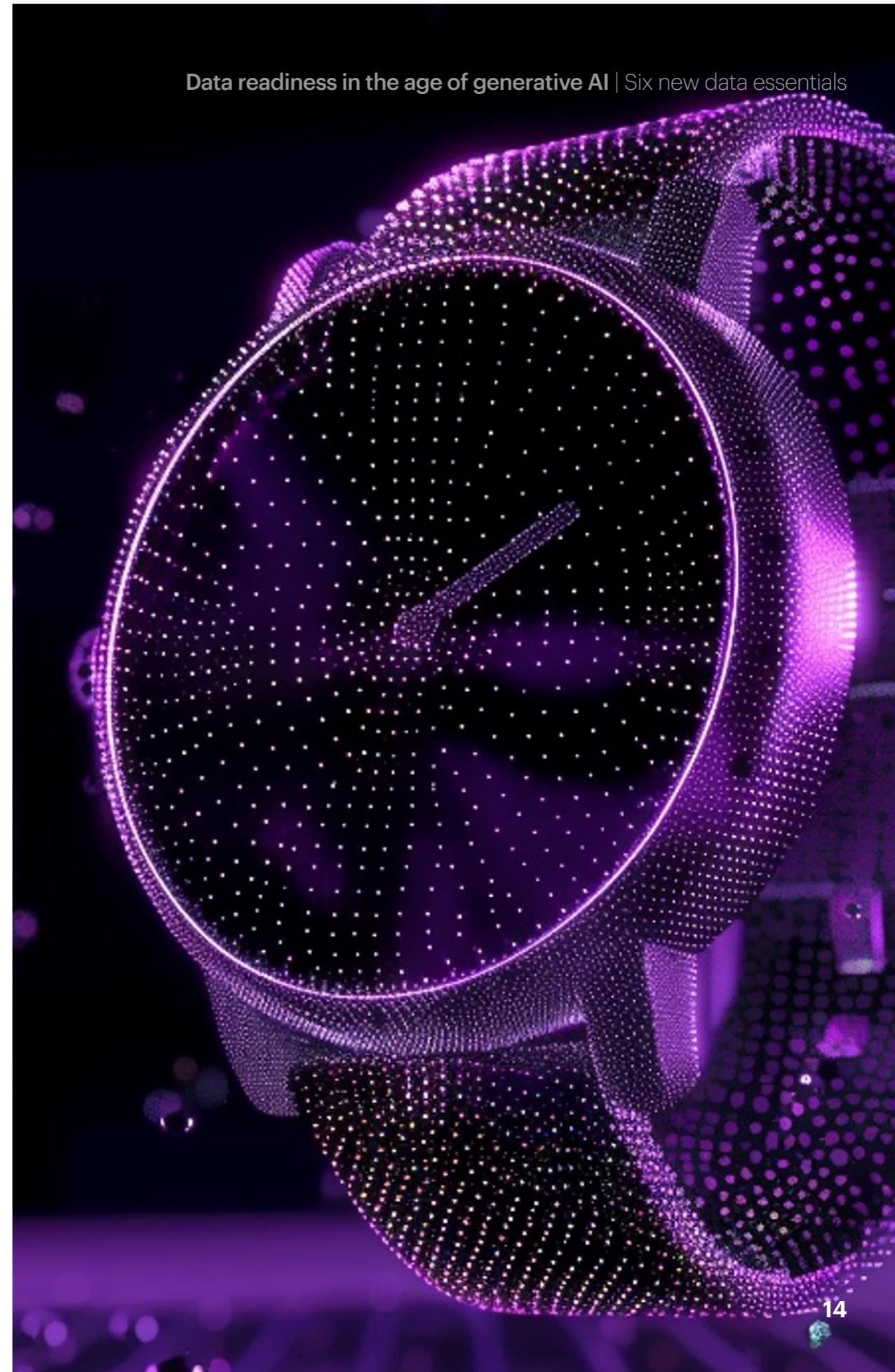
To ensure business accountability and return on investment is front of mind, establish a value framework. This will help you right-size expected outcomes against costs—for example, increased productivity by enabling data scientists and engineers, net new innovation and reduced risk by enabling business users, and even monetization both indirectly with partners and directly with third parties.

Operationalize this framework with observability tools that automatically monitor the quality and consumption of your data, to systematically capture value and pave the way for new investments that maximize your data's potential.

“We have consistently prioritized value quantification in our data and analytics initiatives, so much so that we established a dedicated value workstream. Their analysis has enabled us to confidently attribute a significant return on investment to many initiatives over the last two years. Much of this quantified value stems from strategic business initiatives enabled by our data platforms, including dynamic personalization fueled by customer data, our data-driven loyalty programs and improvements in restaurant operations.”

Matt Sandler

Senior Director of Data & Analytics, McDonald's



2

Reinvent your data architecture and governance to account for new opportunities and risks

Companies that have already invested in a foundation of defined standards for secure, trusted data will be better set up for generative AI experimentation. But there is room for all companies for further industrialization.

Extend your governance and operating model for broader data use across the business and even data sharing with partners. Improve data literacy and awareness about the new data opportunities and risks across your business.

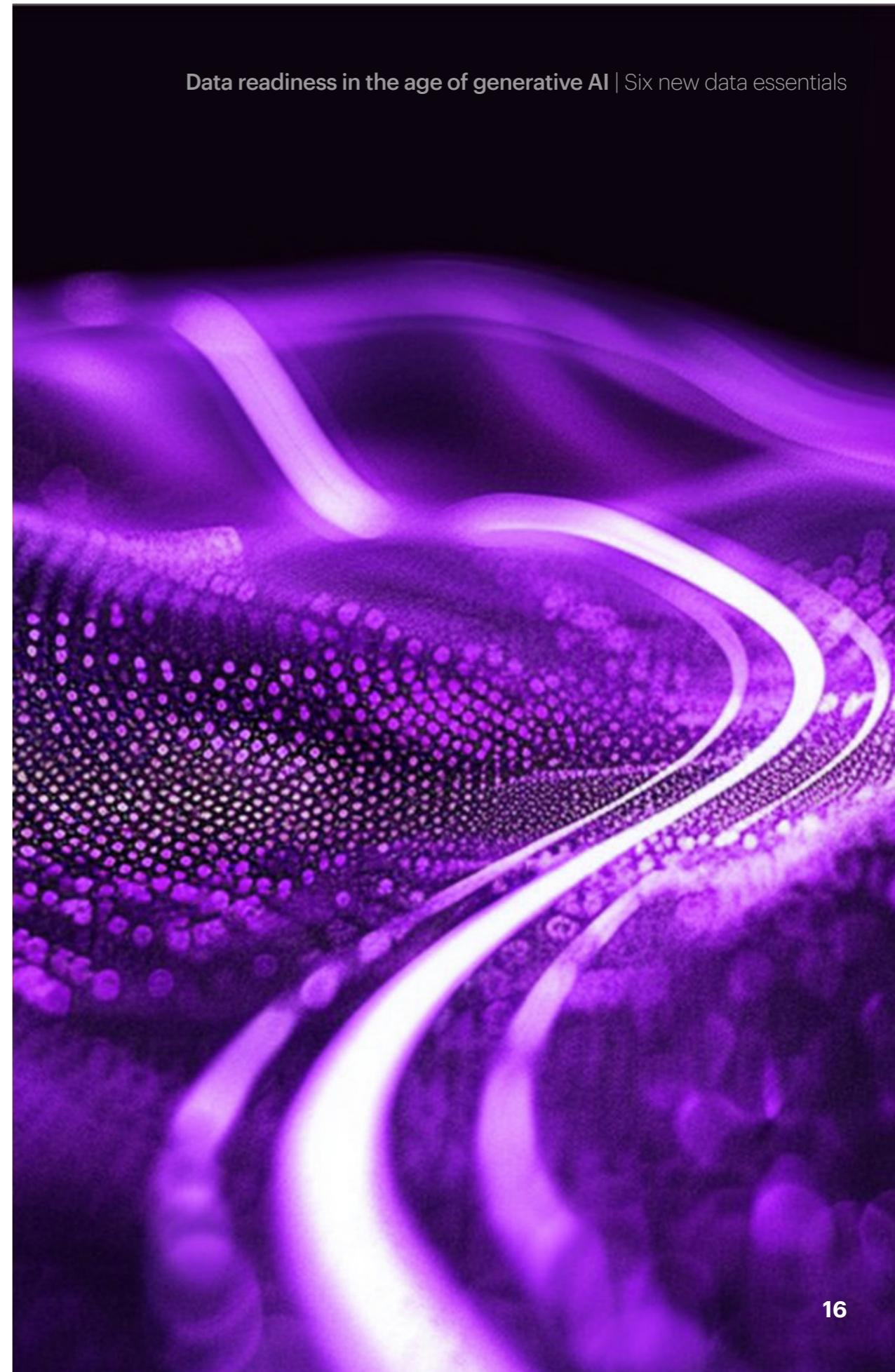
Unlock your proprietary data for by working with ecosystem partners to extend your data architecture for scaled applications of generative AI.



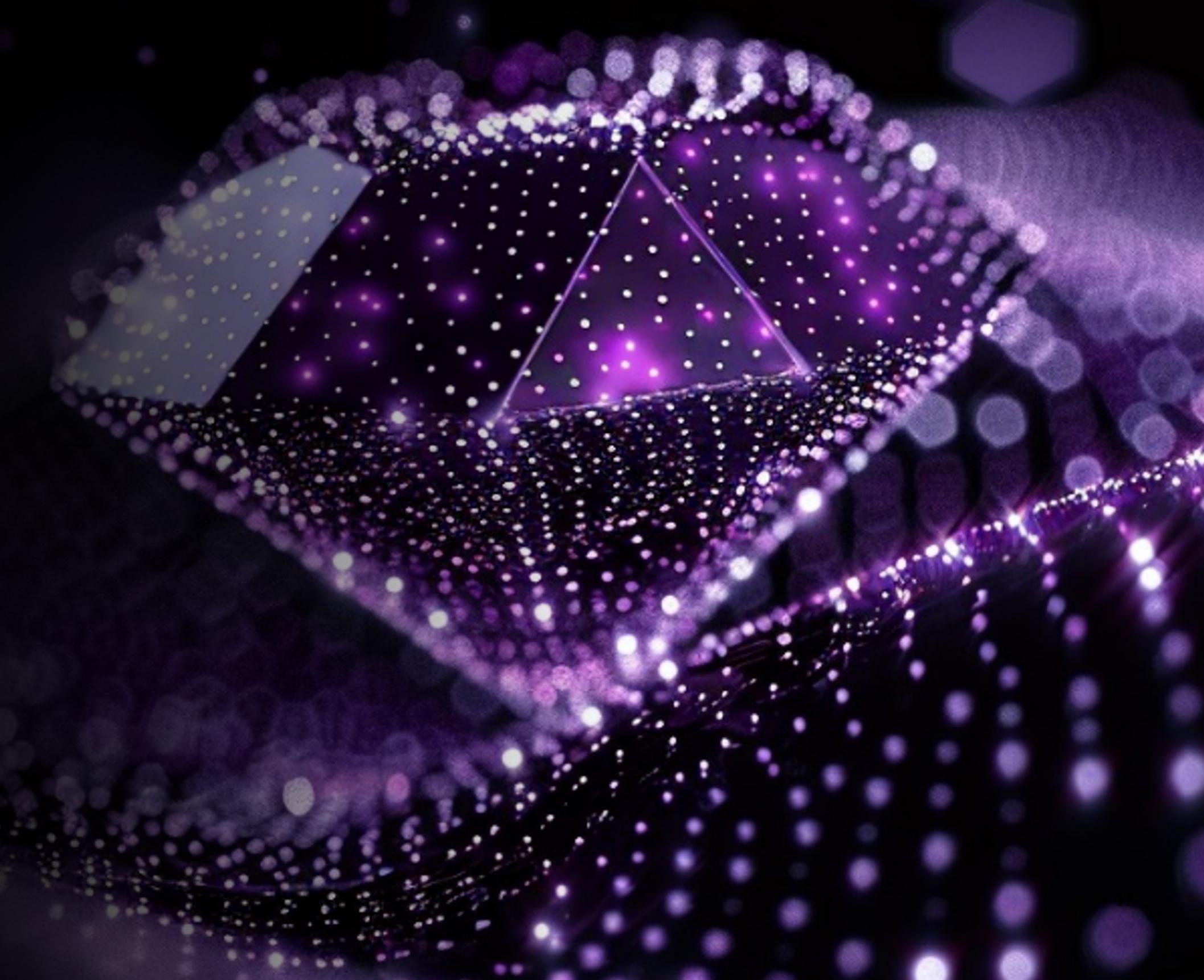
"Every organization wants to be a data and AI company — that's the only way they'll stay ahead of the competition. However, many are still struggling to transition generative AI projects from pilot to production due to privacy, quality, and cost concerns. Databricks Mosaic AI provides support for building and deploying compound AI systems, which offer higher production quality, lower costs and accurate, safe and governed AI applications. We deliver data intelligence: AI that can reason on a company's proprietary data"

Ali Ghodsi

Co-founder and CEO, Databricks



The good news is that the data industry is working together to build the new capabilities needed. For example, building a semantic layer that captures your company's domain expertise, protecting confidential data, and combining structured, synthetic, and unstructured data.



“It’s estimated that between 80% and 90% of the world’s data is unstructured, representing a vast, untapped reservoir of insights. Imagine what your teams could achieve if they could tap into that information—getting summaries of sales meeting transcripts or sentiment trends from customer service calls or social media communications. By harnessing unstructured data, organizations can enhance customer support, pinpoint emerging issues, accelerate operations and maintain a competitive edge. Now, envision taking this a step further: What if your teams could query this data using natural language, asking questions like, ‘What were the top issues in this quarter’s customer service calls?’ These are the kinds of real-world applications we’re already seeing with Cortex AI, showcasing the transformative power of tapping into unstructured data.”

Sridhar Ramaswamy

CEO, Snowflake



3

Apply generative AI to reinvent your data

If you're ready to tap your data's potential, generative AI powered applications and agents can help accelerate your journey, so that you can jump start your data readiness and scale generative AI. For example, accelerating data migration and modernization projects with autonomous agents or creating fewer business intelligence (BI) reports and dashboards and more generative AI enabled interfaces.

This approach leads to Circular Data Pathways: generative AI creating better data products, which then supplies data to all parts of the business – including generative AI models themselves. The data supply chain becomes powered by generative AI to drive new insights and experiences, which, in turn, augments the data supply chain.

“We are leveraging generative AI to revolutionize the creation of metadata for our data products and marketplace, build a semantic layer that serves as the conceptual backbone of our business understanding, and create knowledge graphs that enhance the accuracy, speed and contextual richness of strategic insights. This comprehensive approach has not only improved and accelerated our data journey but also enabled us to rapidly scale our metadata ecosystem with rich, robust and accurate content, thereby driving confidence, user engagement and the ultimate adoption and application of our data assets.”

Kyle Pudenz
SVP, Enterprise Data & Analytics, Cencora



Generative AI is changing data as we knew it – elevating its importance, changing the ecosystem and creating new requirements that companies must address. The six new data essentials are a starting point for companies on the road to data readiness. Now that you've got this information, how will you move forward and ready your data?

The road to data readiness

Remember—you already own the most valuable asset in the era of generative AI: your data. Now you need to make sure your data is ready for generative AI.

By understanding the six data essentials and the key actions that you can take to improve your data readiness, you too can pull ahead of the pack.

Six data essentials in the age of generative AI

01

Your proprietary data is your competitive advantage.

02

Your unstructured data holds untapped potential.

03

Synthetic data is key to filling in data gaps.

04

Connected data is the key to provide context for generative AI.

05

Generative AI accelerates data risks.

06

Generative AI, applied to data, jumpstarts data readiness.

Key actions to improve data readiness

01

Understand the value and invest to maximize your data's full potential.

02

Reinvent your data architecture and governance for new opportunities and risks.

03

Apply generative AI to reinvent your data.

About the research

This report draws on insights from Accenture's work on more than 1,800 client projects, as well as Accenture's "Art of AI Reinvention" survey. The latter—fielded from June to July 2024—covered 2,000 executives (including CEOs, Chief AI Officers, other CXOs and data-science executives) at companies with revenues greater than \$1 billion and located in diverse industries.

The survey aimed to understand how organizations design, develop and deploy AI models to create both financial and non-financial value. Topics covered included: data and AI strategy, data and AI architecture, strategic bets, budgets and investments, talent strategy, ecosystem strategy, responsible AI and data and AI challenges and adoption.

Acknowledgements

Data & AI team

Rahul Basole, Scott Lee, Prateek Peres-da-Silva,
Ekta Sahni, Pragya C Sharma, Mudit Srivastava, Ajay Vasal

Editorial and research team

Sophie Burgess, Francis Hinterman, Praveen Tanguturi

References

1. <https://www.accenture.com/us-en/blogs/cloud/why-global-leaders-think-generative-ai-game-changer>
2. Ibid
3. Art of AI Reinvention survey, 2024
4. <https://www.accenture.com/us-en/insights/technology/data-products>
5. Art of AI Reinvention survey, 2024
6. Ibid
7. <https://www.accenture.com/in-en/insights/artificial-intelligence/synthetic-data-speed-security-scale#>
8. Art of AI Reinvention survey, 2024
9. https://www.engadget.com/the-new-york-times-is-suing-openai-and-microsoft-for-copyright-infringement-181212615.html?_fsig=gEA6xbnejbztE4NPnmvogQ--%7EA
10. Art of AI Reinvention survey, 2024
11. <https://artificialintelligenceact.eu/>
12. <https://newsroom.accenture.com/news/2024/s-p-global-and-accenture-partner-to-enable-customers-and-employees-to-harness-the-full-potential-of-generative-ai#:~:text=S&P%20Global%20and%20Accenture%20Partner%20to>

About us

About Accenture

Accenture is a leading global professional services company that helps the world's leading organizations build their digital core, optimize their operations, accelerate revenue growth and enhance services—creating tangible value at speed and scale. We are a talent and innovation-led company with 774,000 people serving clients in more than 120 countries. Technology is at the core of change today, and we are one of the world's leaders in helping drive that change, with strong ecosystem relationships. We combine our strength in technology and leadership in cloud, data and AI with unmatched industry experience, functional expertise and global delivery capability. Our broad range of services, solutions and assets across Strategy & Consulting, Technology, Operations, Industry X and Song, together with our culture of shared success and commitment to creating 360° value, enable us to help our clients reinvent and build trusted, lasting relationships. We measure our success by the 360° value we create for our clients, each other, our shareholders, partners and communities.

Visit us at www.accenture.com

Copyright © 2024 Accenture. All rights reserved. Accenture and its logo are registered trademarks of Accenture.

This content is provided for general information purposes and is not intended to be used in place of consultation with our professional advisors. This document refers to marks owned by third parties. All such third-party marks are the property of their respective owners. No sponsorship, endorsement or approval of this content by the owners of such marks is intended, expressed or implied.

Authors



Karthik Narain

Group Chief Executive – Technology & Chief Technology Officer



Senthil Ramani

Lead – Data & AI



Lan Guan

Chief AI Officer



Teresa Tung

Lead – Data Capability



Amit Bansal

Lead - Data Markets

