Google Cloud

Data and AI Trends Report 2024

The impact of generative AI

New opportunities, new technologies, new skills.

Gen AI is here – and it's a game-changer! This revolutionary technology will disrupt industries and transform our lives more profoundly than ever before. Data is the fuel for AI, and what powers its effectiveness. To truly take advantage of gen AI in your enterprise, you need the ability to access, manage, and activate your structured and unstructured data across a variety of systems.

Furthermore, your data can also benefit from AI and machine learning (ML) for deeper understanding, to enhance models, or improve customer experiences. Success hinges on achieving all of this while maintaining a high level of data quality and security, while upholding responsible data use principles.

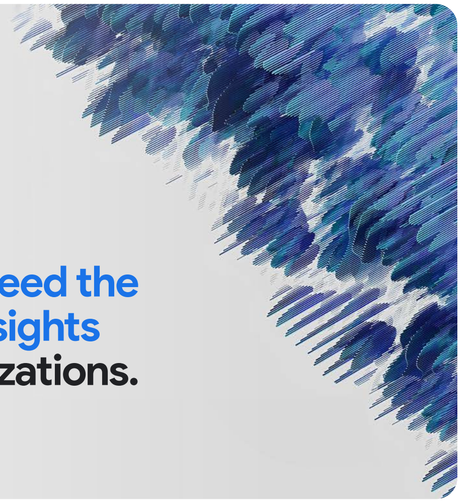
Changes are rippling across the entire data stack in response to this new era. To learn more about how technologies are shifting, Google surveyed hundreds of business and IT leaders with questions about their goals and strategies for harnessing gen AI. This report delves into their perspectives for 2024 and beyond, offering valuable insights for organizations looking to capitalize on gen AI within their enterprise.

Top 5 trends at a glance:

1

Gen AI will speed the delivery of insights across organizations.

5 minute read



2

The roles of data and AI will blur.

6 minute read

3

AI innovation will hinge on strong data governance.

3 minute read

4

Operational data will unlock gen AI potential for enterprise apps.

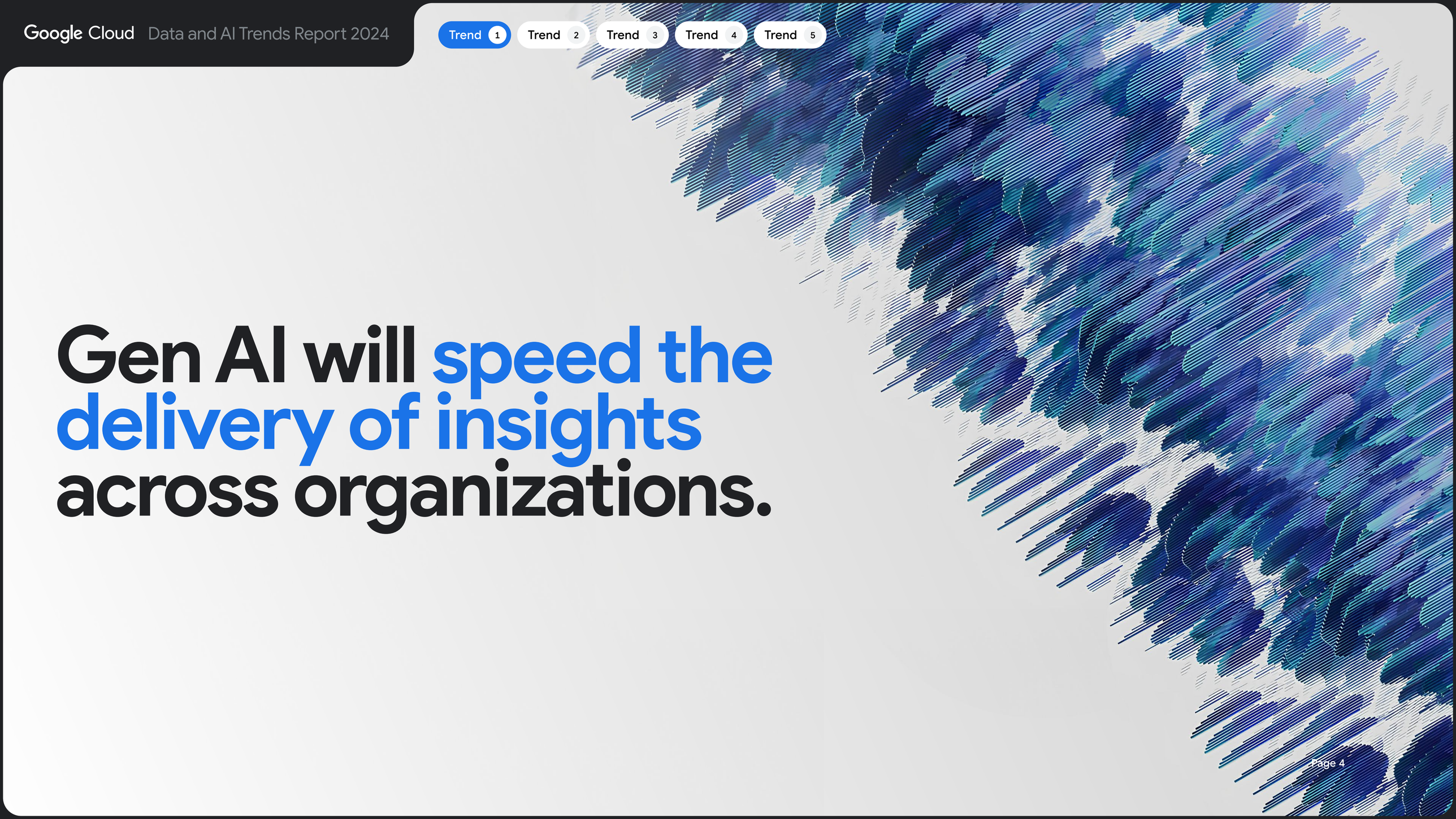
4 minute read

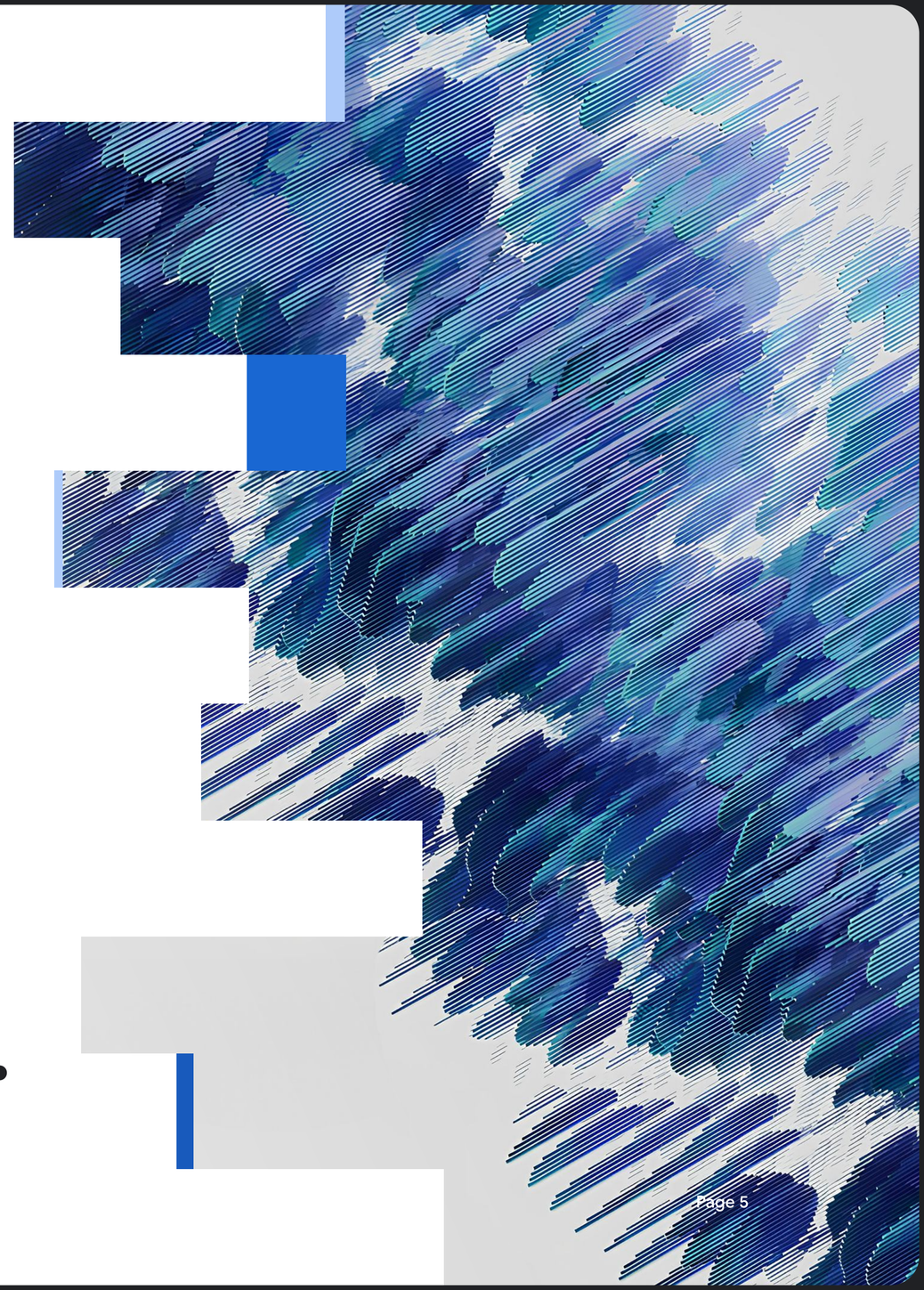
5

2024 will be the year of rapid data platform modernization.

5 minute read

Gen AI will speed the delivery of insights across organizations.





Nearly 2/3 of data decision makers expect a democratization of access to insights in 2024.

84% believe gen AI will help their organization access insights faster.

Google Cloud Customer Intelligence Trends Research Survey, 2024.

Why should you care?

It's almost impossible to overstate how significantly gen AI has changed the technological landscape. In the case of business intelligence (BI), as tools become more accessible, even non-technical team members will be able to benefit from these insights; driving productivity and disseminating knowledge faster than ever before. That means better data literacy across your organization, smarter decisions being made, and ultimately greater success in the market.

52% of non-technical users are already using gen AI to draw out insights today.

Modern BI tools were already developing ways to bring data to everyone who needed it; reports embedded in the most relevant context for the data, such as account insights appearing in a salesperson's CRM, is an easy example. But those insights have always needed to be carefully curated by an analyst. The end user has always been a step removed from the data. Connecting a large language model to your business data closes that gap. Team members can interact with your data intuitively and conversationally, or create reports and dashboards by simply 'talking' to your data or making a simple search across your business. In fact, many of the organizations surveyed for this report are already putting this into practice.

AI is already used by both the most advanced data scientists and within lines of business.

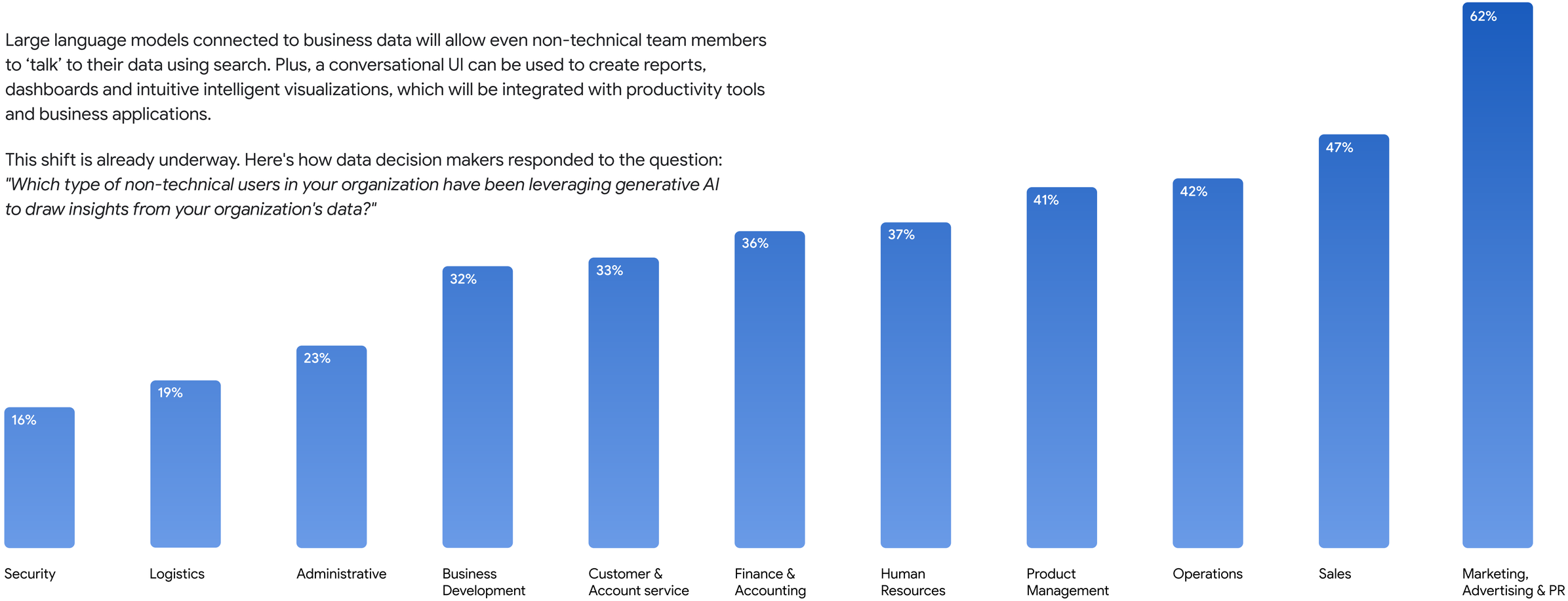
Essentially, tools that connect people to key business data through natural language will be a major force in bridging existing gaps in organizational skill sets.

Throughout 2024 and beyond, expect to see more business users ‘talking’ to their data using search and leveraging a conversational UI to create reports, dashboards and intuitive intelligent visualizations, which will be integrated with productivity tools and business applications. As many applications allow users to see how others found successful answers to questions, people will also be able to benefit from aggregate knowledge, as well as gaining insight into which interactions have had the greatest impact over a day, a quarter, or a year.

AI for all.

Large language models connected to business data will allow even non-technical team members to ‘talk’ to their data using search. Plus, a conversational UI can be used to create reports, dashboards and intuitive intelligent visualizations, which will be integrated with productivity tools and business applications.

This shift is already underway. Here's how data decision makers responded to the question: *"Which type of non-technical users in your organization have been leveraging generative AI to draw insights from your organization's data?"*



Google Cloud Customer Intelligence Trends Research Survey, 2024.

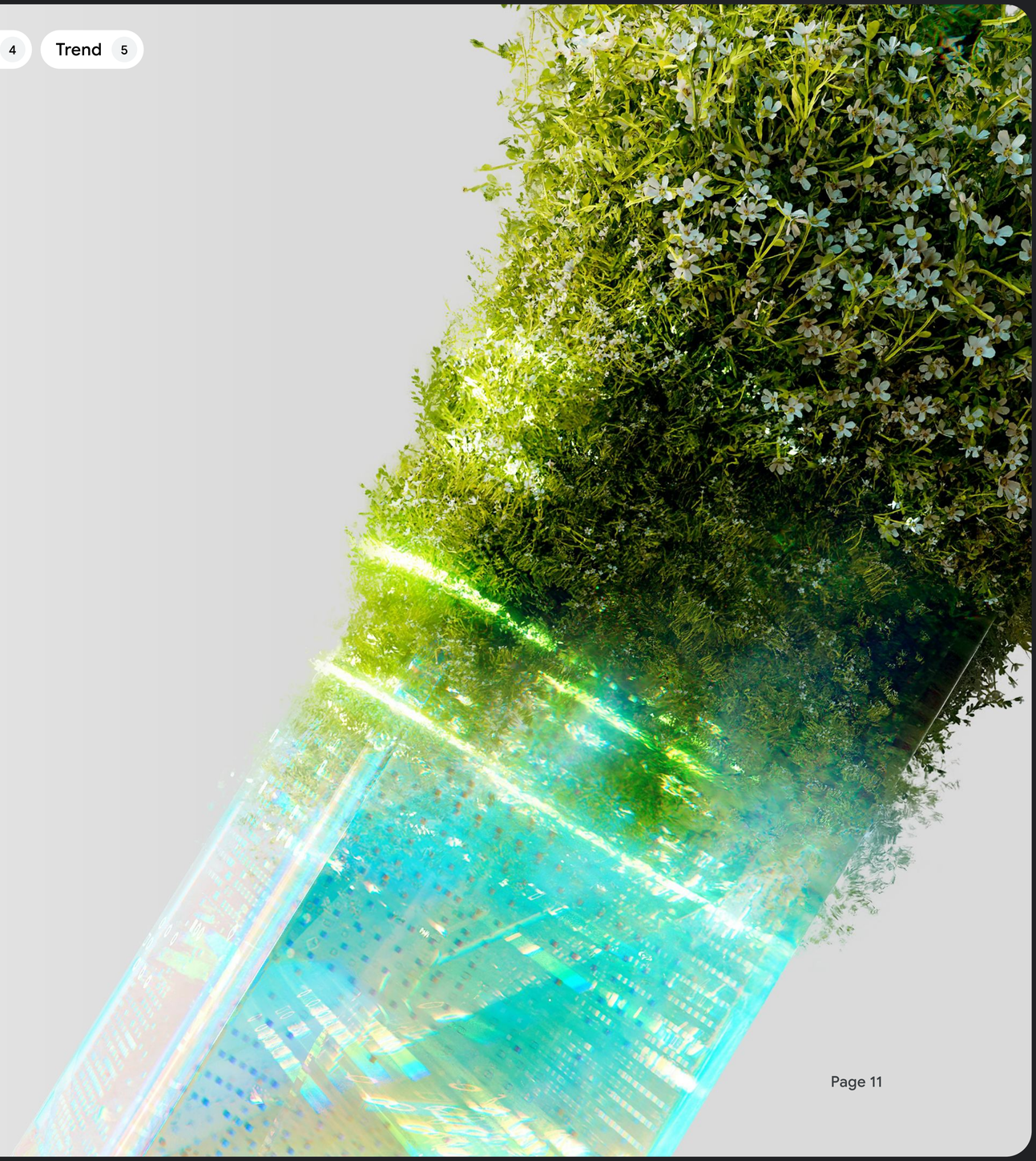
"Wendy's introduced the first modern pick-up window in the industry more than 50 years ago, and we're thrilled to continue our work with Google Cloud to bring a new wave of innovation to the drive-thru experience. Google Cloud's gen AI technology creates a huge opportunity for us to deliver a truly differentiated, faster, and frictionless experience for our customers, and allows our employees to continue focusing on making great food and building relationships with fans that keep them coming back time and again."

TODD PENEGOR
President and CEO, Wendy's

“Gen AI isn’t just enhancing BI, it’s reinventing it. In 2024, insights won’t be uncovered – they’ll be proactively surfaced, with more nuance and context than ever before.”

IRFAN KHAN
President & Chief Product Officer, SAP HANA Database & Analytics, SAP

The roles of data and AI will blur.



80% of respondents agree that the lines between data roles are starting to blur.

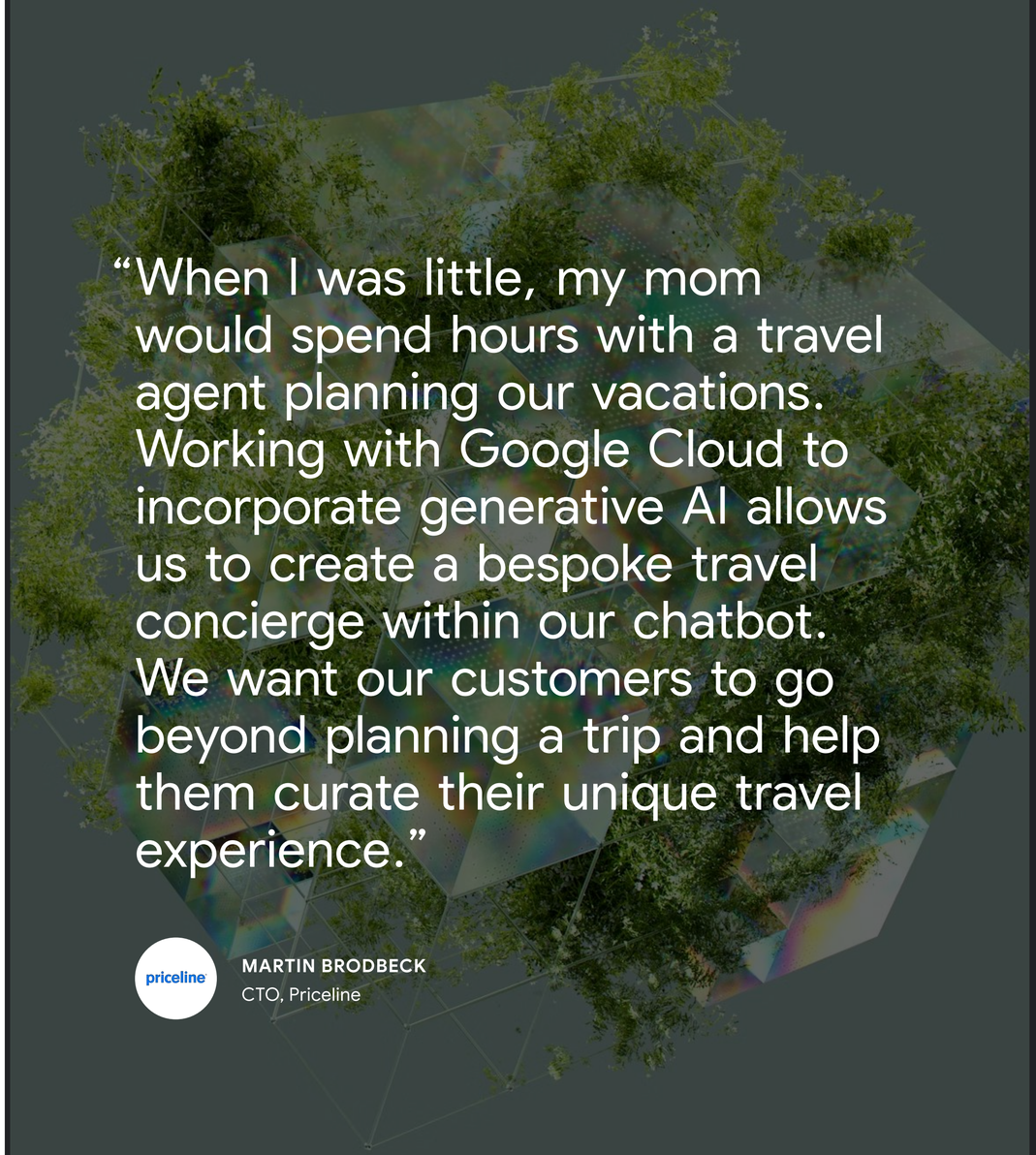
Google Cloud Customer Intelligence Trends Research Survey, 2024.

Why should you care?

As the use of AI becomes more widespread, the speed at which companies can go from raw data to AI will become increasingly important.

The organizations that master this process will be able to make better decisions, launch new products and services faster, and provide superior customer experiences.

As processes are streamlined, the roles of data and AI will become increasingly blurred; meaning formerly siloed teams will need to work more closely together than ever before.



Many data analysts are now taking on responsibilities that were traditionally reserved for data scientists and vice versa.

Data analytics and engineering, AI, and business analytics are the most scarce skills within organizations. Gen AI presents an opportunity to boost productivity of existing data teams and workloads, thus assisting with this widening skills gap. To be able to seamlessly use data and AI platforms allows organizations to improve productivity, and innovate faster by accelerating their data to AI journey.

Data and AI tools are also becoming increasingly interconnected in order to help users streamline data and AI workflows.

This interlocking allows users to:

Have a common workspace for data engineers, analysts and scientists that supports multiple coding languages such as SQL, Python, and Spark.

Extend software development best practices such as CI/CD, version history and source control to data assets, enabling better collaboration and hand-offs.

According to research, more than half (54%) of digital leaders say **skills shortages prevent them from keeping up with the pace of change.**

Nash Squared Digital Leadership Report, 2023.

Gen AI is also providing employees with ways to accomplish more technical tasks. **For instance, tools can suggest the new lines of code required to update a financial-reporting system or outline the A and B versions of a marketing campaign** or otherwise create first drafts that human employees can take and implement into live production environments.

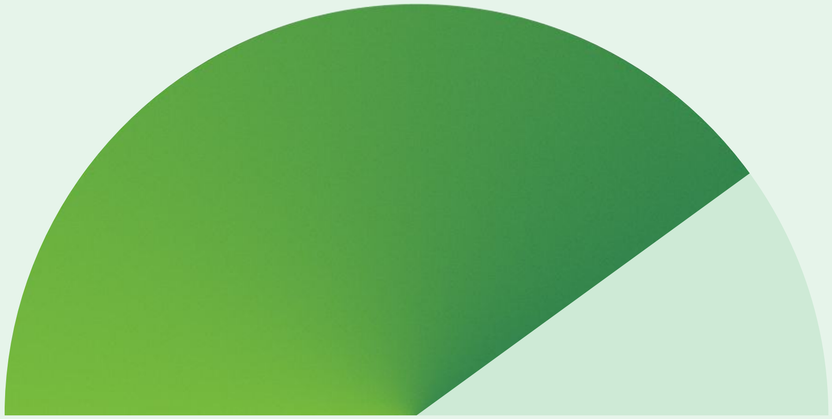
The organization of the future: Enabled by gen AI, driven by people, McKinsey & Company, 2023.

Bringing AI directly to data can activate its full potential regardless of its format.

A significant challenge hindering organizations from fully utilizing the potential of data lies in the substantial amount of untapped, unstructured data generated today. This includes formats such as images, documents, and videos. It is estimated to cover roughly up to 80% of all data, which has so far remained untapped by organizations.

Structured data, characterized by its organization in fixed fields and columns, such as in spreadsheets or databases, can be easily processed and analyzed using traditional methods. However, unstructured data - think social media posts, emails, customer call recordings, clinical documentation, and sensor readings - is often complex and challenging to interpret, making it difficult to extract meaningful insights.

The advent of advanced AI and Machine Learning (ML) technologies has revolutionized the way organizations leverage their data. These cutting-edge technologies offer unparalleled opportunities to unlock the full potential of all data, regardless of its format; structured, semi-structured, or unstructured. Similarly, multi-modal AI has opened up a world of possibilities for organizations, unlocking new levels of efficiency and accuracy when tuning and grounding models in their enterprise data. Text embeddings enable vector searches directly on data, without the need for complex and time-consuming preprocessing steps. This simplifies the process of finding relevant information, identifying patterns and trends, and clustering similar unstructured data in sources like documents.



80% of the global datasphere will be **unstructured by 2025.**

VentureBeat, 2022.

“The most transformative aspect of 2024's data and AI landscape isn't just about efficiency – it's about democratization. By seamlessly interconnecting these technologies, we empower not just data scientists, but business users across the organization to unlock actionable insights and drive innovation.”

ALI GOLSHAN
CEO, Gretel

AI innovation will hinge on strong data governance.

66% of organizations have at least half of their data dark, posing significant risk.

Five Factors For Planning A Data Governance Strategy, Forbes, 2023 & Gartner Glossary, Dark Data, 2024.

Why should you care?

This explosion of new technology has its drawbacks, too. Many organizations are discovering new vulnerabilities and weaknesses, especially when it comes to the quality of their data. It's not enough to just apply LLMs to data – these models need to be grounded in good quality enterprise data or otherwise risk hallucinations. Organizations which take a practical approach to data governance, quality, and trust will be in a strong position to deliver tangible business outcomes with AI.

Most respondents are only somewhat confident (45%) in their organization's data quality, and another 11% are even less than somewhat confident.

Similarly, most respondents (54%) consider their organizations only somewhat mature when it comes to data governance and only 27% consider their organizations either extremely or very mature.

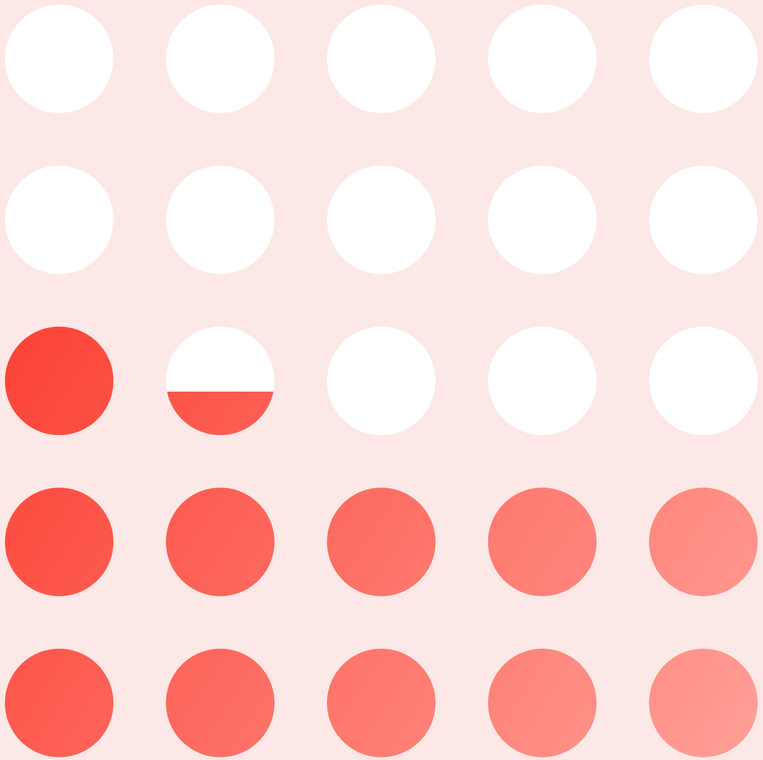
On the plus side, many organizations are already taking steps to ensure data accuracy, data quality, and trust. The majority of organizations surveyed:

- believe they are building a data driven culture.
- are centralizing data governance oversight.
- are building centralized policy management, monitoring, and auditing.

Google Cloud Customer Intelligence Trends Research Survey, 2024.

Fewer than half of respondents (44%) are fully confident in their organization's data quality.

Google Cloud Customer Intelligence Trends Research Survey, 2024.



Snap Inc uses Google's Data Cloud to deliver a business domain-specific, self-service data platform across distributed data, with decentralized data ownership but centralized governance and visibility. With increased data efficiency they can focus on improving the user experience and boosting engagement.

Carrefour uses Google's Data Cloud to achieve zero trust network protections, improving data security and strengthening secure access to business-critical applications. Their data-centric infrastructure provides flexibility to make changes very quickly and deliver the highest quality service to their customers.

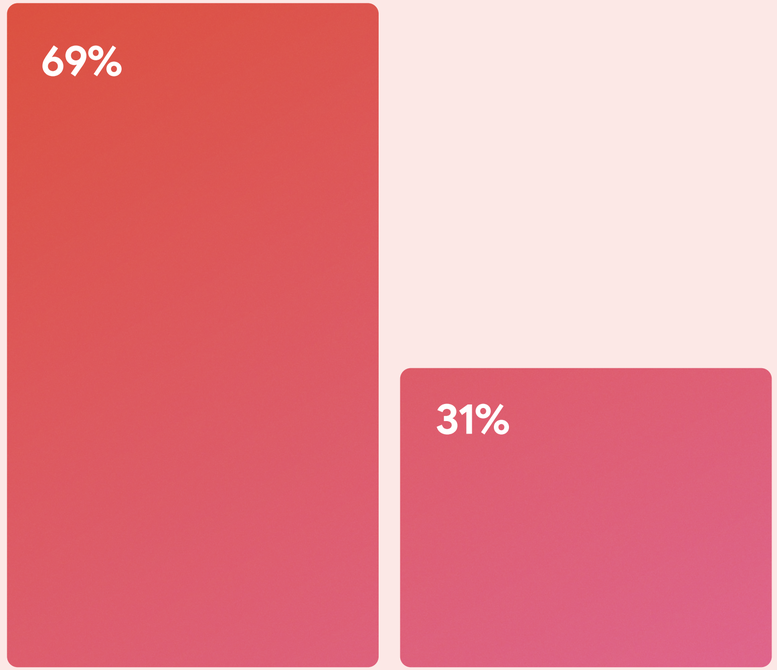
What should organizations look for?

It's key that organizations look for secure-by-design data platforms that fully integrate data encryption. The right platform should automatically catalog the data you own and give you the ability to logically unify and organize your data leveraging metadata. This enables you to centrally secure and govern data, based on your business context, and use built-in automation and intelligence around data profiling, quality, lineage, and more to better manage data at scale. This enables:

Data quality. Auto-generate data quality rules to measure for completeness, accuracy, and validity of your data.

End-to-end data lineage. Automatically generated lineage to track data flows, perform impact analysis, and use lineage as a foundation for governance and compliance across data and AI models.

Unified governance for data and AI assets. Central policy management, monitoring, and auditing for data authorization, retention, and classification.



69% of employees had **bypassed their organization's cybersecurity guidance** in the past 12 months.

Gartner Predicts Nearly Half of Cybersecurity Leaders Will Change Jobs by 2025, 2023.

Collibra

“2024 is a watershed moment for generative AI. Organizations that fail to manage the risks throughout the AI development lifecycle will be left behind. Those who proactively establish strong AI governance practices are the ones who will unlock the true potential of this technology.”

FELIX VAN DE MAELE
CEO and Co- Founder, Collibra

Operational data will unlock gen AI potential for enterprise apps.

71% of organizations plan to use databases integrated with gen AI capabilities.

Google Cloud Customer Intelligence Trends Research Survey, 2024.

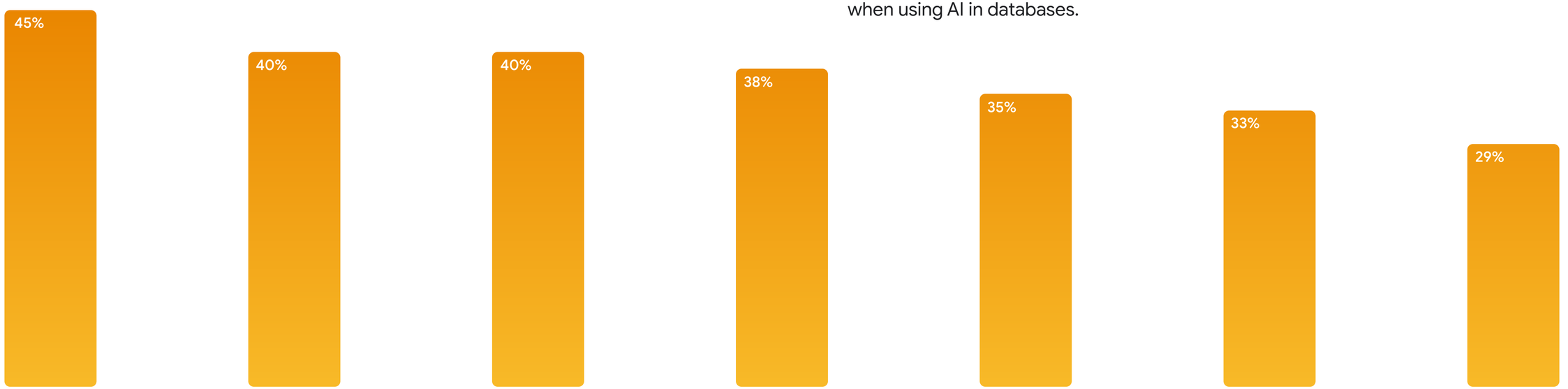
Why should you care?

Businesses are excited about the potential with large language models (LLMs). They have all experienced the power of tools like Gemini and other large language models, but they also recognize that the creative nature of these tools is not a good fit for most enterprise use-cases. Enterprise gen AI applications face a variety of challenges that LLMs alone do not address; they need to provide accurate and up-to-date information, offer contextual user experiences and do all this while not breaking the bank.

Operational databases and warehouses with vector support help bridge the gap between LLMs and enterprise gen AI apps. This is why we're seeing so much interest in vector search and vector databases and why Retrieval-Augmented Generation (RAG) is an important technique for enhancing and augmenting LLMs and gen AI models. We're seeing a lot of innovation across the industry and much of it is driven by the open source community including PostgreSQL, one of the most popular databases for developers.

What do organizations want from AI-powered databases?

Seamless connectivity to AI models, the ability to ground LLMs using techniques like RAG and the ability to use natural language for database administration are the most important capabilities when using AI in databases.



Google Cloud Customer Intelligence Trends Research Survey, 2024.

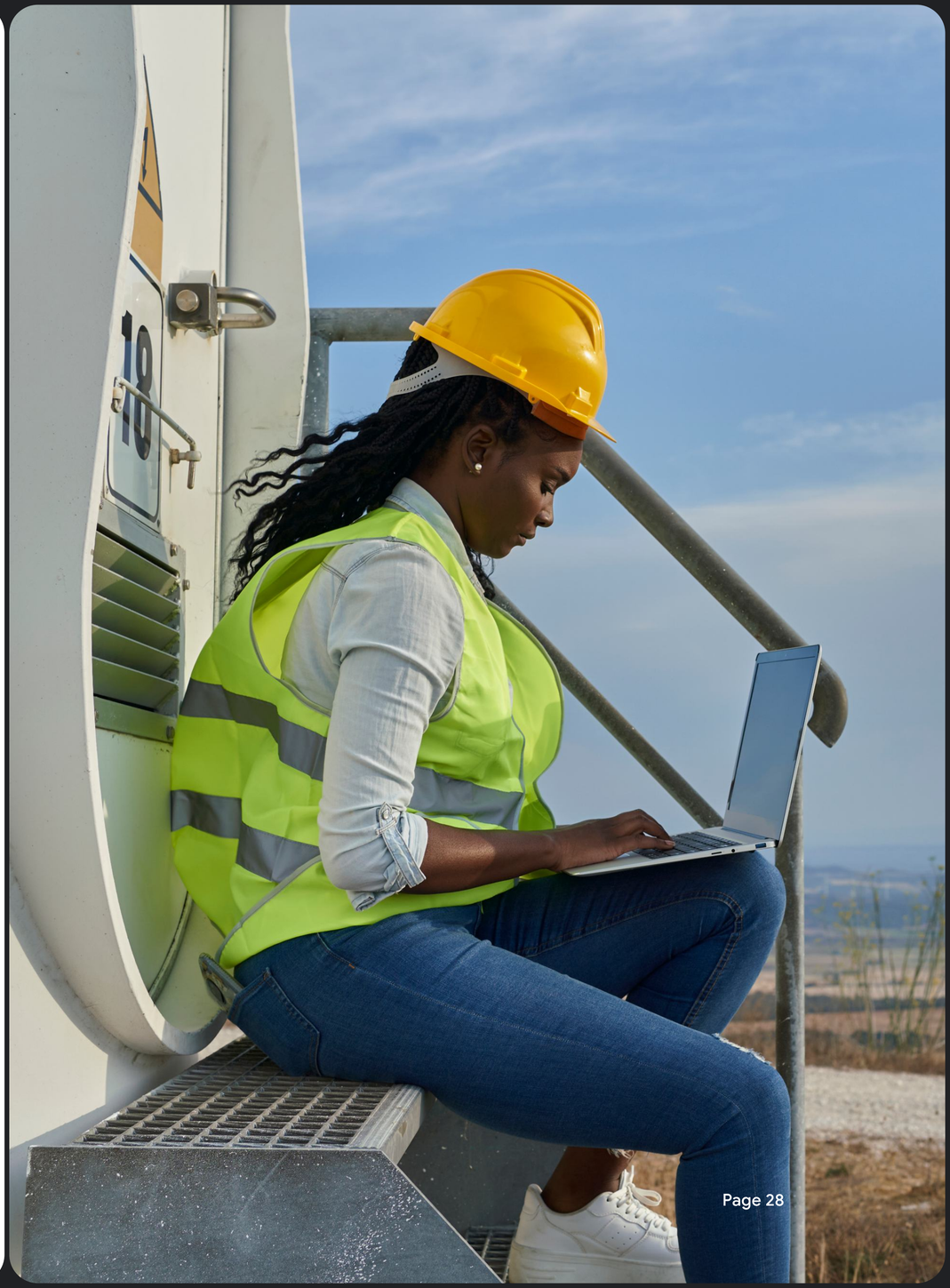
Get it right and reap the rewards.

The true power of gen AI is unlocked when operational data is integrated with gen AI to deliver real-time, hyper-personalized, and contextually-relevant experiences across enterprise applications. Simply put, gen AI-enabled operational databases holding relevant business data will be the key to unlocking gen AI in the enterprise.

Successful databases will evolve to be AI-first, and deeply integrate technologies such as; vector search, seamless connectivity to AI models, support for natural language to SQL, and tight integrations with AI tooling and open source frameworks. All these will be natively built into operational databases and will become table stakes.

Databases that fail to integrate gen AI capabilities are likely to become obsolete.

Having AI closer to the operational data will also allow developers to iterate quickly and enhance the experience with all available data. You can do this where your data already lives because databases are already powering all applications, so organizations don't have to learn or set up an entirely new system and it is significantly more cost effective. In addition, with open source technologies like PostgreSQL, developers can get started quickly with familiar tools and capabilities.



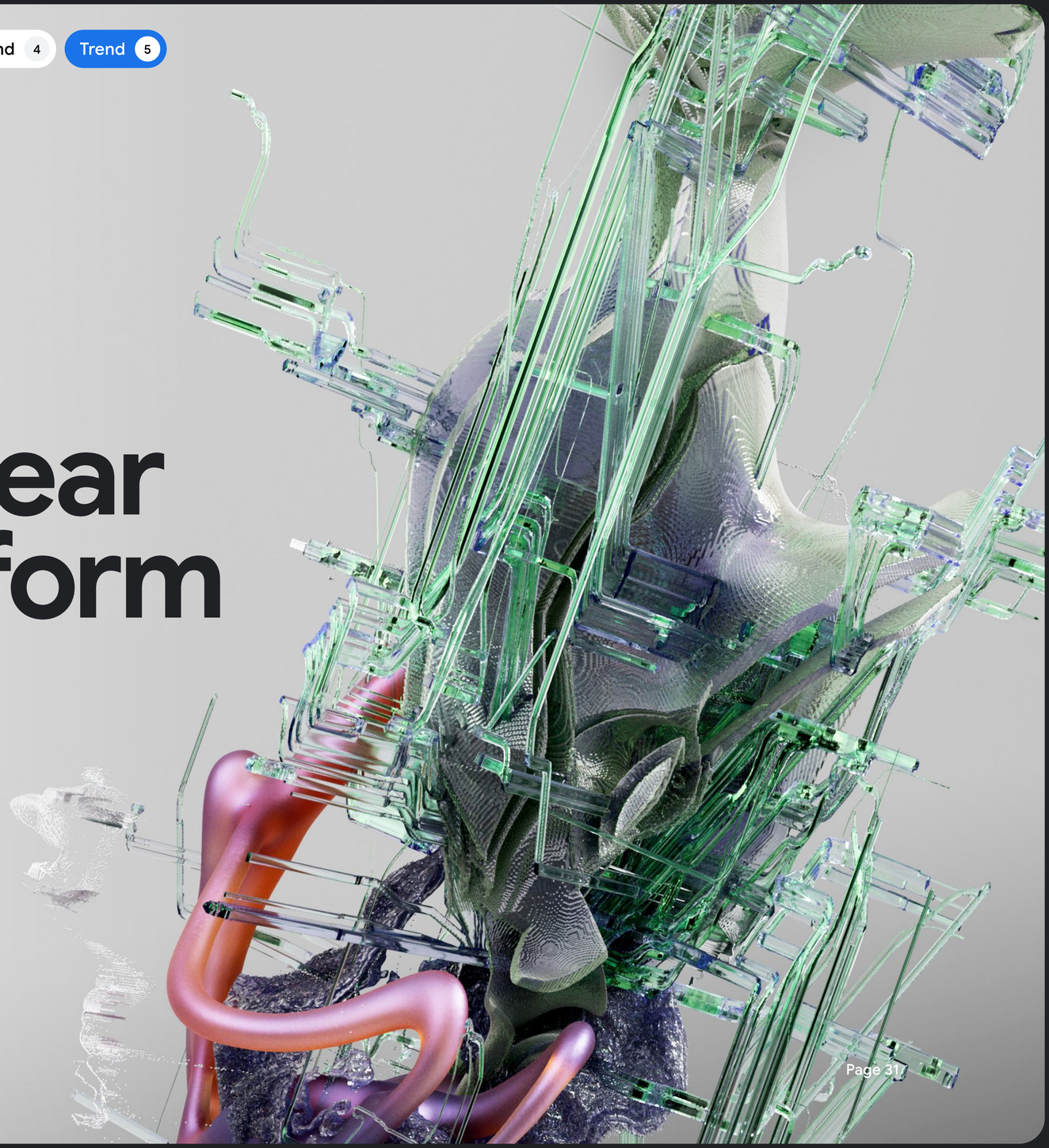
“We explored several new entrants in the database market that focus on storing vectors and ended up trialing several. And given Linear’s existing data volume and our goals for finding a cost-efficient solution, we opted for Cloud SQL for PostgreSQL once support for pgvector was added. We were impressed by its scalability and reliability. This choice was also compatible with our existing database usage, models, ORM, etc. This meant the learning curve was non-existent for our team.”

TOM MOOR
Head of US Engineering, Linear

“Customers are looking to leverage the power of LLMs by augmenting them with their domain knowledge and enterprise data. To support these new use cases, cloud-based database solutions that also embrace open-source gen AI orchestration frameworks will provide application developers with the capabilities to help them quickly and more efficiently build Retrieval Augmented Generation (RAG) applications.”

HARRISON CHASE
CEO, Langchain

2024 will be the year of rapid data platform modernization.



Only 14% of organizations are satisfied with their legacy databases' support for AI, indicating there is a lot of room for improvement.

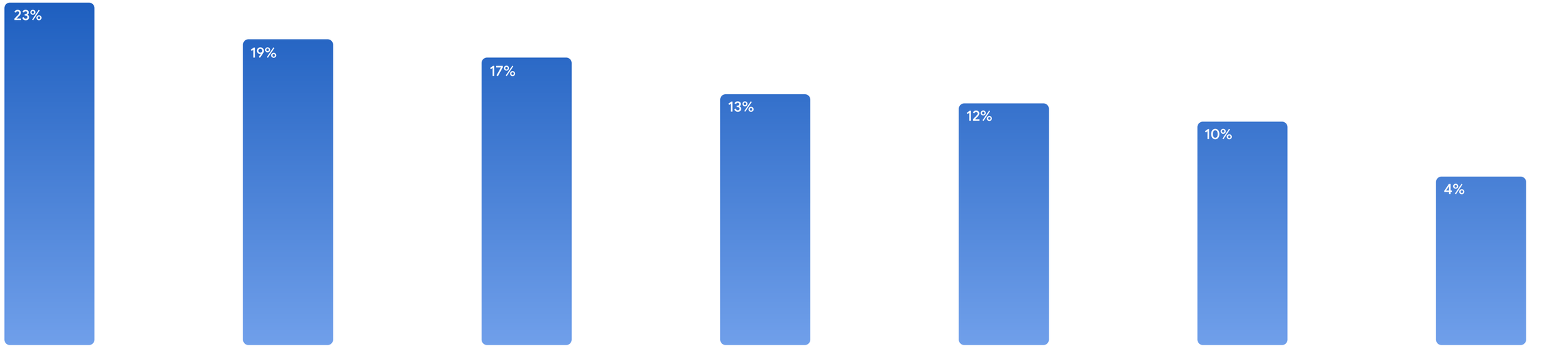
Google Cloud Customer Intelligence Trends Research Survey, 2024.

Why should you care?

As more and more organizations seek to take advantage of the opportunities gen AI brings, many are discovering their legacy databases are holding them back due to lagging technology and poor user experience. In addition to outdated technology and a poor developer experience, legacy databases have also caught the attention of C-level executives because of their expensive, unfriendly licensing and vendor lock-in, which often result in millions of dollars in unnecessary annual costs.

The gen AI boom is bringing new urgency to database modernization because the most popular AI tools for working with vectors, models, and data run in the cloud and are based on open source database technologies such as PostgreSQL. In addition, the most advanced AI models run only on major cloud platforms.

Top challenges with legacy databases.

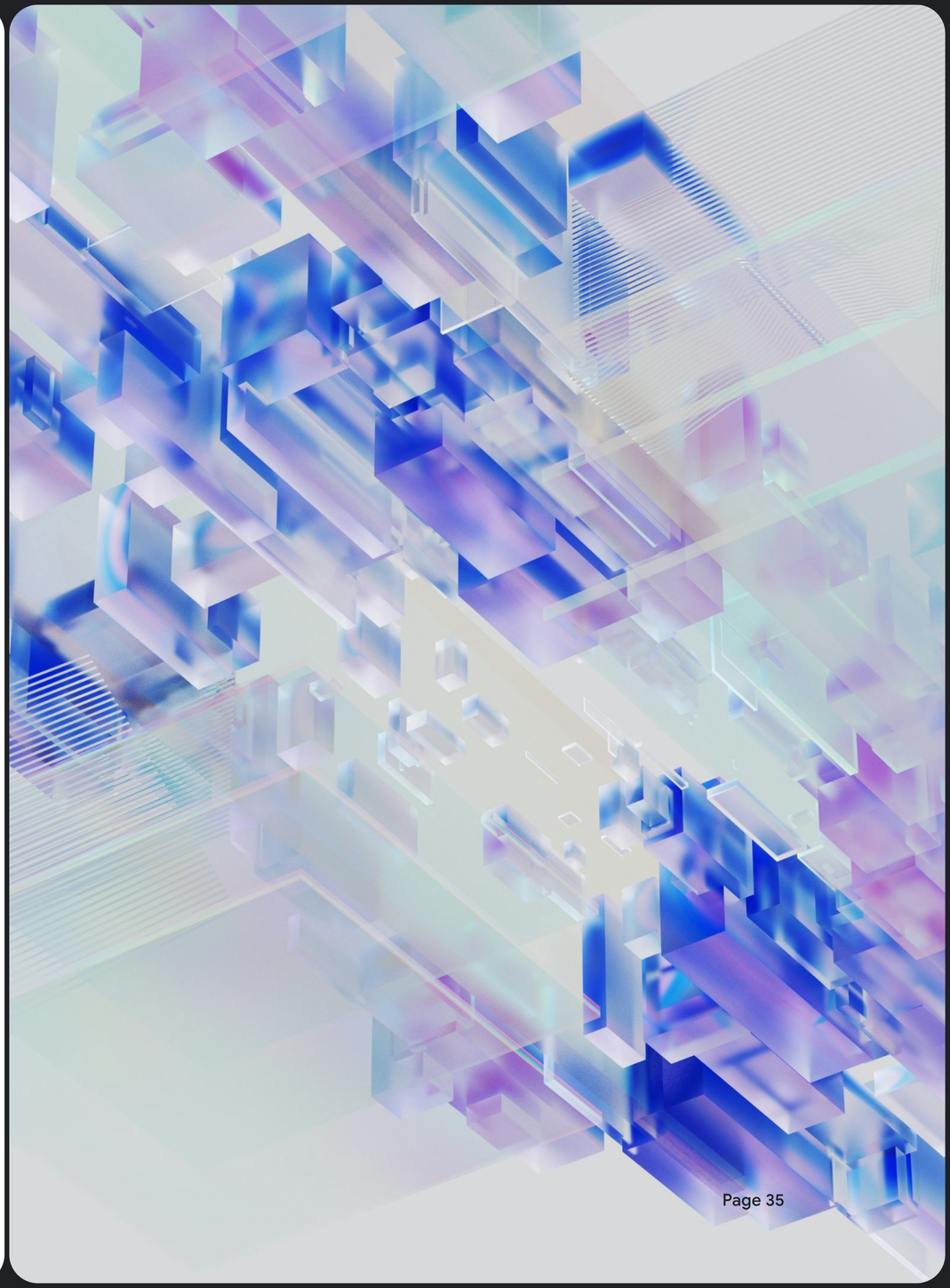


Google Cloud Customer Intelligence Trends Research Survey, 2024.

Smooth transitions are more possible than ever.

Thankfully, migrating from legacy databases is becoming easier with database migration tools and programs continuing to improve and mature. We're also seeing AI help to augment these tools, to the point where breaking free from legacy databases is now much easier with AI-assisted code conversion, code completion, and improved efficiencies.

IT decision makers are now comfortable approving large modernization projects as they look to embrace open technologies, including gen AI, as part of their innovation roadmaps.



“Data migration tools have been around forever, but more recently, they've been getting smarter with the ability to do AI assisted code conversion and code completion. The hardest part of the migration is transforming the data and training the new applications to fit in the new database. Both are made easier through gen AI. You can use a model to look at a source database and find out how to transform the data into the destination database. You can get some quick wins, and ultimately get developer productivity. There's still a ton of legacy stuff, and gen AI is bringing the bar down to simplifying migrations.”

ANDREW STORRS
VP Data Engineering, Aritzia

“Character.AI is a pioneer in the design and development of open-ended conversational applications. Our gen AI platform utilizes our own advanced neural language model to generate human-like text responses and engage in contextually relevant conversations. When we found AlloyDB for PostgreSQL, we were stuck between a rock and a hard place. Usage of our service had scaled exponentially, putting unique stresses to various parts of our infrastructure, especially our databases. Google Cloud's AlloyDB and Spanner provide a solid foundation, delivering reliability, scalability, and price performance for our workloads, from engagement and operations, to AI and analytics.”

JAMES GROENEVELD
Research Engineer, Character.AI

“Legacy systems were built for a different era. Gen AI's potential lies in its ability to find novel connections and insights. In 2024, organizations that liberate their data from outdated systems will be best positioned to stay ahead of the curve.”

GOPAL SRINIVASAN
Principal - Generative AI, Deloitte Consulting

An entire generation of developers is building AI applications and leveraging AI for more efficient coding, improved database performance insights, and enhanced security posture. Are you one of them?

How Google Cloud can help.

Google Cloud helps organizations unify data and connect it with groundbreaking AI to unleash transformative insights and personalized experiences. By harnessing the simplicity, scalability, security, and intelligence of Google's unified data and AI approach, businesses can unlock the full potential of their data in a single, streamlined solution.

Because Google Data Cloud consolidates workloads and manages the entire data life cycle, data teams are empowered to develop modern, data-driven applications using popular open-source engines and models.

We also provide cutting-edge AI/ML and generative AI capabilities that are readily available for your data, enabling all of your people to easily and quickly access the data they need and unlock its true value. All of this is delivered with enterprise-grade efficiency. It’s this unique combination that makes Google Cloud an unparalleled partner for turning raw data into organizational value.

Inside our one-of-a-kind approach.

Fully connected data and AI.

As new ways of interacting with systems and data emerge, it's clear that organizations need AI models grounded in quality enterprise data that allows for analytical insights and augmented experiences.

With Google's Data Cloud, data teams can use gen AI tools to activate their enterprise data across BigQuery and AlloyDB, and use built-in features to easily apply AI/ML directly to their data. For instance, BigQuery ML allows data teams to construct ML models straight on their BigQuery data simply using SQL, and even call foundation models in Vertex AI. Built-in vector embedding capabilities in AlloyDB also allow users to store and generate embeddings within their data stores to help augment their LLMs and support their gen AI use cases.

A unified data foundation.

Google Cloud's unified data foundation is built on BigQuery, and brings your data together into one place, integrating structured and unstructured data with AI to deliver insights across your entire data estate. This unified data foundation allows you to manage your entire data lifecycle and help make data access, management, governance, and analysis easier for different types of users within an organization, effectively removing data silos.

Our highly scalable architecture unifies transactional and analytical systems, enabling tightly integrated data services across BigQuery, AlloyDB, and Spanner. This allows easy data analysis from Spanner to BigQuery, with virtually no impact to the underlying transactional workloads.

The most open data platform for modernization.

Google Cloud is committed to being the most open cloud provider, letting you build modern, data-driven applications wherever your workloads are. We support open source and open standards, and offer managed database services that are fully compatible with popular open-source engines and models.

With AlloyDB Omni and BigQuery Omni, you can utilize data and modernize your applications across Google Cloud, AWS, Azure, and Google Distributed Cloud, without incurring the costs, security risks, and governance concerns associated with data migration. It's now easier than ever to get started with gen AI on a data platform that meets you where you are on your modernization journey.

Enterprise-grade efficiency and security at scale.

Google Data Cloud is an industry leader in efficiency, security, and scale; catering to organizations of all sizes and adhering to the most stringent enterprise requirements.

We help make it easy for organizations to share data safely and securely across organizational boundaries, run queries across exabytes of data with blazing speed, and process billions of transactions – all with generally lower cost.

So, what's next?

Clearly, 2024 is shaping up to be an exciting - and pivotal - year for many organizations. Those who are able to prepare their people and platforms to fully embrace new capabilities made possible with gen AI will not only see short-term productivity gains, but begin to effectively future-proof their organization against ever-evolving competition.

Naturally, many of these new opportunities require new skills, and for existing processes to be refined. Organizations that embrace the need to upskill and fully equip their people will quickly find that this investment pays dividends in the form of almost limitless potential.

Ready to join the party?

If you've got any questions about the content of this report, or want to know more about how Google Cloud can support your organization, our experts are always on hand.

Talk to an expert

You can also take our **Data & AI Strategy Assessment** to discover how ready your organization is for AI-powered digital transformation and receive expert recommendations to get you there faster.

Take the assessment

Methodology.

The Google Cloud Customer Intelligence team conducted a global research study on Data & AI Trends with 410 Data Decision Makers from 12/18/2023 - 1/17/2024. Active recruitment was paused from 12/23 to 1/1 for winter break. Respondents included a mix of Data, IT, and business leadership roles with seniority ranging from C-level to Manager. All respondents were employed at 1,000+ employee organizations currently using data products & services. Respondents did not know Google was the research sponsor and the identity of participants was not revealed to Google.

Region

NORAM	180
EMEA	104
JAPAC	76
LATAM	50

Industry

Financial services	74
Retail (e.g. Grocers, Stores, Boutiques, Franchises, Restaurants, etc.)	42
Technology	85
Other	103

Company size

1,000 to 4,999	19%
5,000 to 9,999	20%
10,000 to 49,999	30%
50,000+	31%

Role level

C-level	14%
VP or equivalent	20%
Director	43%
Manager	7%
Lead / Head	4%

Interaction with data products and services

Hands on	40%
Strategic/oversight	60%

Role

Business Development	2%
IT or IS (Information Technology, Computer Engineering, Security, etc.)	40%
Software Development	1%
Technology Strategy or Product Development	11%
Marketing/Advertising/PR	16%
Operations	3%
Product Management	3%
Research/Analytics/Strategic Planning	6%
Sales	1%
Data Science	15%

Google Cloud

Data and AI Trends Report 2024