CLOUDERA

WHITEPAPER

Survey Report

The State of Enterprise AI and Modern Data Architecture

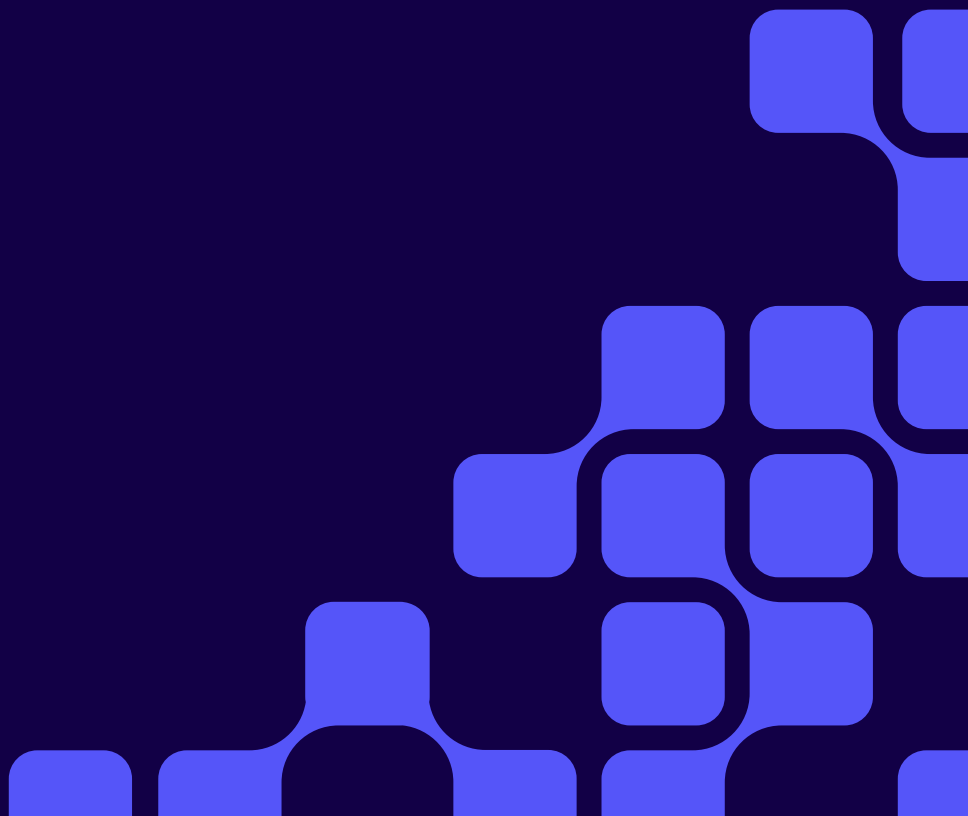
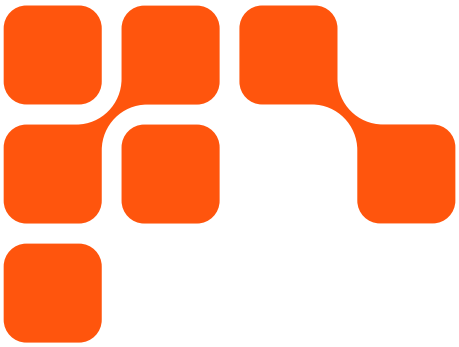


Table of Contents



Introduction

Enterprise Artificial Intelligence (AI) and the vast potential it holds are dominating discussions across industries. As organizations push to modernize, boost efficiency, and generate greater business value, Enterprise AI has emerged as a key enabler for those goals to be achieved in new areas of business operations.

But is the potential of AI being realized? Are organizations utilizing AI apps in production to make business decisions or are they still experimenting? As it turns out, despite the widespread interest in Enterprise AI, evidence suggests that transformative value remains elusive for many companies primarily due to the limitations of the outdated data infrastructures that are powering AI tools.

To underscore this point — and to educate audiences on the positive impact AI already generates — Cloudera, in partnership with third-party research firm Researchscape, conducted a survey of 600 IT leaders based in the U.S., EMEA, and APAC regions on:

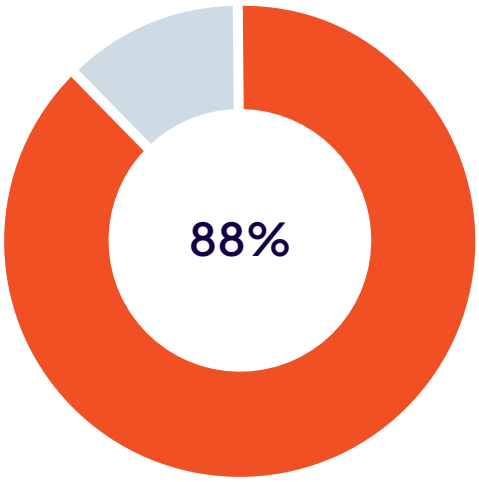
- The challenges and barriers with Enterprise AI adoption across global businesses
- Current applications or deployment plans for Enterprise AI across the organization
- The state of data infrastructures — e.g., where data is stored, who has access, and more
- The benefits of hybrid data management in relation to Enterprise AI adoption

The Current State of AI: Why

In a sea of possibilities, enterprises are implementing AI applications at a rapid pace, with organizations in Italy (98%), Spain (96%), Brazil (95%), France (87%), and the United States (88%) seeing the greatest adoption.

It's clear that businesses have wasted no time in putting AI to work within their organizations as 88% of survey respondents said their company is currently using AI. And for good reason: this technology boasts the ability to completely transform the way businesses operate and is primed to revolutionize industries. The top three departments that are leveraging AI are IT (92%), Customer Service (52%), and Marketing (45%). Whether it's automating and streamlining IT processes, building a chatbot capable of supporting customer needs quickly and effectively, or leveraging analytics to foster better decision-making, AI is being integrated into nearly every facet of the business.

The way AI is utilized within those departments (IT, Customer Service, and Marketing) manifests in a few different ways. As the technology continues to permeate various industries, IT leaders are prioritizing improvements to both internal and external operations. The survey found that the top three benefits that organizations experienced from leveraging AI are improved customer experiences (60%), increased operational efficiency (57%), and expedited analytics (51%). Companies are leveraging AI within their organizations to completely redefine or elevate customer experiences by applying the technology to enhance security and fraud detection (59%), automate aspects of customer support (58%), leverage predictive customer service (57%), and power chatbots (55%).



of survey respondents said their company is currently using AI.

Top 5 Reasons Why Organizations are Looking to Adopt and Scale AI:

(Ranked as #1 reason among respondents)

- Increasing operational efficiency (21%)
- Driving innovation (17%)
- Staying competitive (11%)
- Improving customer experiences (10%)
- Enhanced decision making (9%)

Figure 1

The Current State of AI: How

AI comes in all sorts, shapes, and sizes. For our purposes, we'll define three of the most commonly deployed kinds of AI — Generative AI (GenAI), predictive, and deep learning.

GenAI encompasses tools like large language models (LLMs), natural language processing (NLP), and image generators, and it is primarily focused on creating new content and data based on learned patterns.

Next, as one may infer from its name, predictive AI is all about forecasting what might happen in the future. This form of AI utilizes machine learning (ML) algorithms to predict future possibilities based on patterns that can be identified within historical data. From there, predictive AI will prescribe actions to help capitalize on those opportunities.

Finally, deep learning is a subset of ML involving neural networks with multiple layers. It is what powers both generative and predictive models, enabling them to handle complex tasks by learning from large amounts of data.

Of all the AI implementations out there, GenAI has cemented itself firmly as the most popular deployment today. Sixty-seven percent of respondents said that they are running generative models in some form.

Companies are using a variety of architectures when employing GenAI applications, including: exposing data to open-source models that do not get trained on their data (60%), exposing data to open-source models that become trained on their data (57%), bringing open-source models to data where foundational models are trained on premises or in a private cloud (50%), and building proprietary LLMs or Small Language Models (26%).

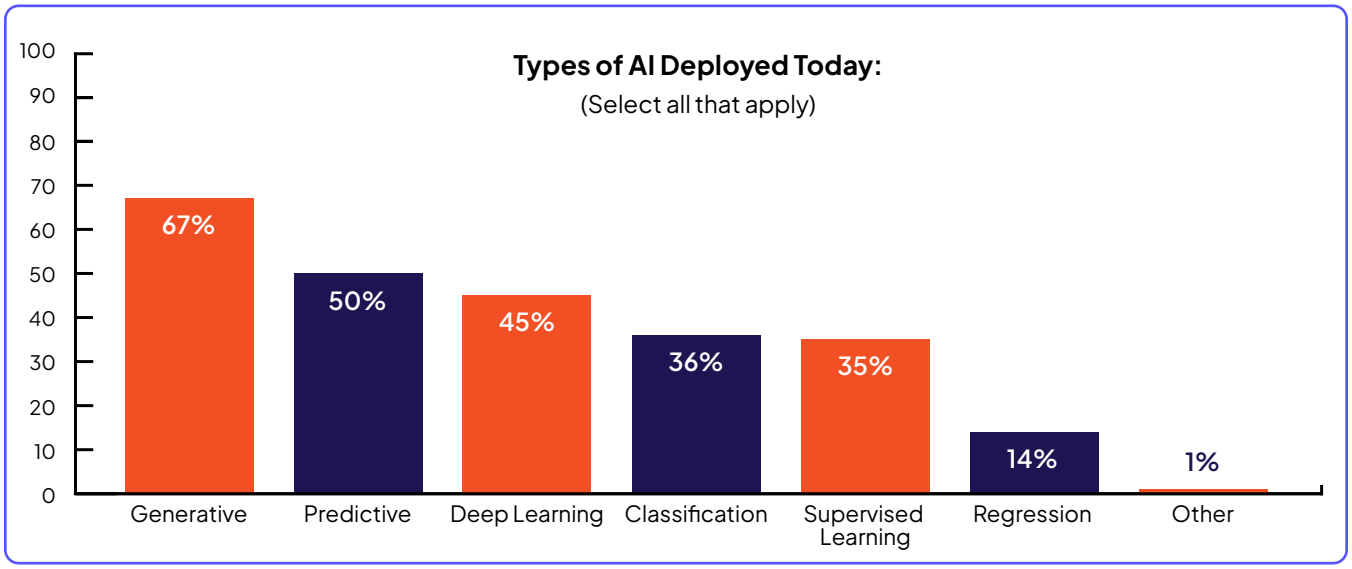


Figure 2

Challenges to Enterprise AI Adoption

AI adoption may be widespread among organizations, but when it comes to the scale of adoption, enterprises are less aggressive. More than a third (39%) said that only some or almost none of their employees are using AI currently.

As companies accelerate their AI initiatives, cost factors continue to be a challenge. The top three cost prohibitions for accelerating AI identified by respondents are integration costs (50%), data storage costs (49%), and cost associated with data breaches or leakages (46%).

An important part of tackling these challenges is to ensure there is a strong base of talent in place within an organization. In fact, 38% of respondents noted that they lack the proper training or talent to manage AI tools. As businesses look to solve the myriad of challenges that exist with AI adoption, they are making investments in a number of places, including:

- In-house training and development programs (57%)
- Cross-functional AI teams and projects (57%)
- Hiring and talent acquisition strategies (50%)
- Participation in open source and AI communities (44%)
- Collaboration with academic institutions (30%)

Top three cost prohibitions for accelerating AI:

Integration costs (50%)

Data storage costs (49%)

Cost associated with data breaches or leakages (46%)

Figure 3

Top barriers to adopting AI:

74%

Worried about the security and compliance risks that AI presents

38%

Don't have the proper training or talent to manage AI tools

26%

Cost of AI tools is too expensive

25%

Don't have the data infrastructure required to effectively power AI

8%

Cost to access computer capacity training models is too high

Today's Data Architecture

The challenges most organizations face in trying to implement AI largely stem from the limited or outdated data infrastructures that they rely on to power AI tools. Today, every business has adopted some sort of data architecture, though the kind of architecture they use varies.

A majority of organizations are storing data on a private cloud (81%), but other forms of data architecture also remain popular. These include public cloud (58%), on premises mainframes (42%), on premises distributed (31%), in other physical environments (29%), or in a data lakehouse (19%).

Managing the complexity of today's data landscape presents a unique set of complications. This includes data security and reliability (66%), growing data management costs (48%), compliance and governance issues (38%), overly complex processes (37%), siloed, difficult-to-access data (36%), mistrust in connecting private data and inaccuracies in AI models (32%), and standardized data formats (29%). It's no surprise that of the top three challenges, two are specifically centered around security, compliance, and governance. As threats evolve and new regulations like the Digital Operational Resilience Act (DORA) emerge, any lapse in security or compliance could prove disastrous and costly.

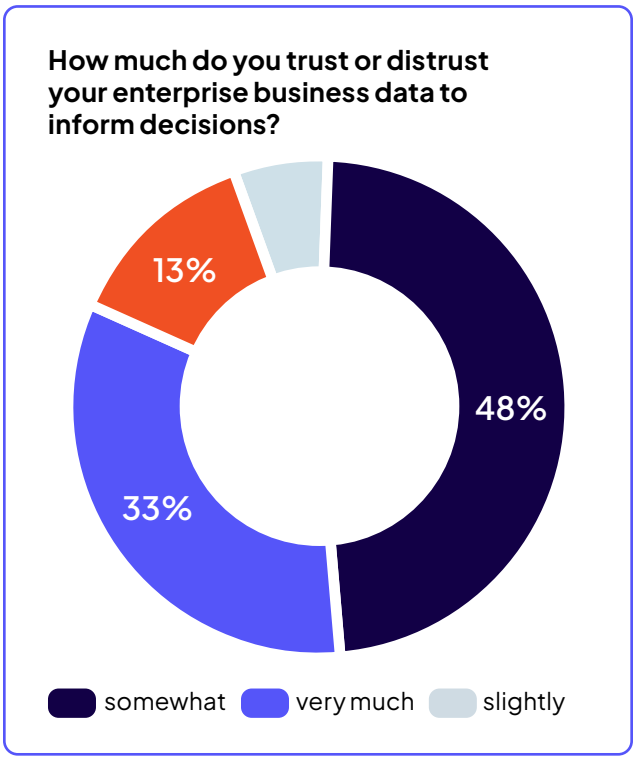
To keep pace, businesses need trustworthy data they can rely on to stay agile and make the right decisions. The top three challenges that organizations face to trusting their data are redundant, contradictory datasets (49%), an inability to govern data across platforms (36%), and having too much data (35%).

Attitudes Toward Data

Even as organizations grapple with challenges surrounding data architecture, employees seem highly confident in their data. In fact, 94% of respondents indicated that they trust their data — though to what degree varies. (See figure 4)

Nearly three-quarters (71%) of respondents say it is "very" or "completely" true that they are 'confident their current data landscape is ready to support Enterprise AI,' and they 'know where their data comes from and are confident in its integrity.' Further, 77% of respondents say it is "very"' or "completely" true that 'their company is using data to get products to market faster.' Another 77% of respondents said it is either "completely" or "very" true that 'their company is using all of the data at its disposal to make smarter business decisions.'

Figure 4



But, digging deeper, only 17% of respondents report that their organization is extremely data-driven, meaning all decisions are based on data analysis and metrics. Likewise, 39% say they considered their organization to be very data-driven, indicating that data analysis plays a significant role in employees' decision-making, though experience and intuition are still relied upon to some extent.

The discrepancy between the number of respondents that believe their data landscape is ready to support Enterprise AI and those that consider themselves extremely data driven presents a disconnect.

As outlined in Figure 5, very real challenges permeate data management within today's data architectures. More than two thirds (73%) of respondents cite that their company's data exists in silos and is disconnected, and while 40% believe they are the sole person who knows where data exists in the organization. How can a company be data driven if their data is disparate and not easily accessible? And if that's the case, how can they be ready for Enterprise AI?

Roughly half (55%) of respondents agreed to some extent that they would rather get a root canal than try to access all of their company's data.

Figure 5

Challenges organizations experience when it comes to data management:

73%

At least some data exists in silos across their organization and is not connected

55%

They would rather get a root canal than try to access all of their company's data

40%

Believe they are the only ones within their organization that knows where data exists

25%

Any employee in the organization can access any data, at any time

Solving the Challenges of Modern Data Architectures

No matter how trustworthy or AI-ready organizations think their data is, the only way to ensure its trustworthiness is by adopting modern data architectures. A modern data architecture provides critical flexibility and visibility for businesses and serves as a blueprint for accelerating the process of gathering insights and value from data. It simplifies data access across organizations and unlocks data from silos, making it easier to understand and act upon.

Modern data architectures also deliver key functionality in terms of increasing the scalability of data management. This architecture can handle data in all forms — structured, semistructured, unstructured — blending capabilities from data warehouses and data lakes into data lakehouses. The industry's best platforms also incorporate functionalities like data processing, data integration, and data governance as part of the overall structure of data systems.

When asked what hybrid data architecture advantages are most valuable to their organization, respondents emphasized data security (71%) as a major benefit. This is followed by improved data analytics (59%), improved data management (58%), scalability (53%), cost efficiency (52%), flexibility (51%), and compliance (37%).

The security, visibility, accessibility, and insights that modern data architectures deliver feed directly into a business's ability to adapt quickly, handle change, and make well-informed decisions — all of which play an important part in future-proofing data infrastructure for the long term and delivering real ROI with AI deployments.

Figure 6

Of companies currently leveraging a data lakehouse, the top benefits are:

33%

Improved operational efficiency

22%

Improved access to unstructured and semistructured data

16%

Reduced data management costs

17%

Strengthened data governance and compliance

11%

Breaking down siloed data

The State of Enterprise AI: A Country-by-Country Breakdown

Across geographies, businesses show some similarities in how they are storing their data. Brazil (93%), the U.S. (83%), France (76%), Italy (74%), and Spain (71%) survey respondents identified private clouds as their top method for storing data. This is followed closely by public clouds, which is the second most popular response across all these countries, with 43% in Brazil, 56% in France, 59% in the U.S., 62% in Spain, and 71% in Italy.

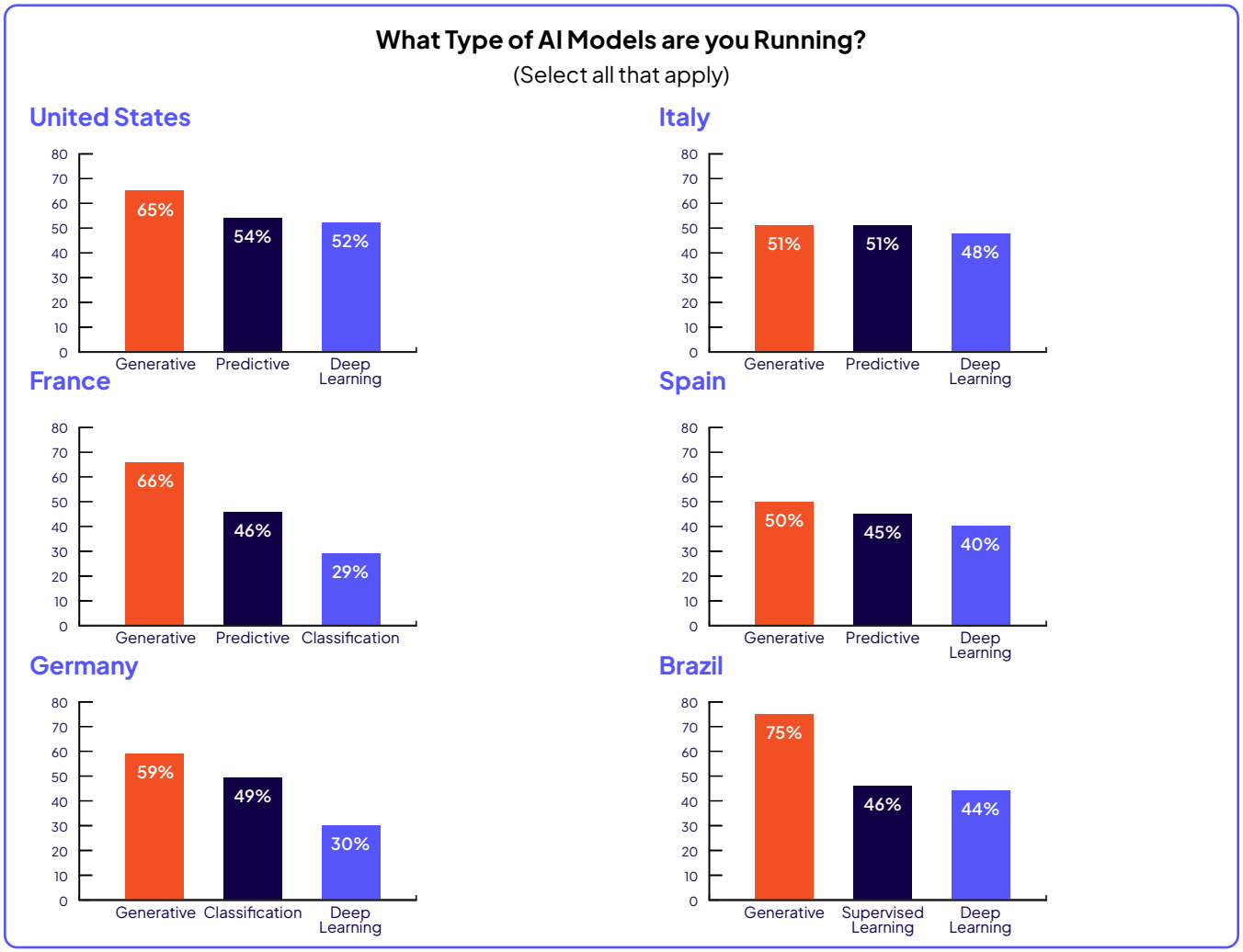
Similarly, there is a consistent trend across regions with regard to how data-driven respondents felt their organizations are — though, to what extent did vary by country. Respondents in Brazil had the strongest viewpoint, with 40% saying their organization is 'extremely data-driven,' and another 44% categorizing their businesses as 'very data-driven.' Comparatively, in France, 43% of respondents said they were 'very data-driven' while only 25% said they were 'extremely data-driven.'

When it comes to the kinds of AI that are being implemented, GenAI remains a clear favorite among businesses, regardless of country. GenAI ranked as the top form of AI used in Brazil (75%), Singapore (70%), France (66%), the U.S. (65%), Germany (59%), and Spain (50%), while it was tied as the top choice with predictive AI (51%) in Italy.

As businesses grapple with the challenges of implementing AI effectively, the issue of trust is always a top priority. As stated earlier, a majority of companies believe their data landscape is ready to support Enterprise AI (71%), but that number jumps considerably higher in select countries, with 91% of respondents in Brazil saying this is completely or very true as did 86% in Italy, and 89% in Spain.

GenAI remains a clear favorite among businesses when it comes to the kinds of AI that are being implemented, regardless of country.

Figure 7



Building Future-Proof Enterprise AI

AI has become a global phenomenon, dramatically altering business operations and accelerating innovation across industries. Regardless of the business, the value that AI can generate is abundantly clear–from delivering better customer experiences to creating more productive, satisfied employees.

But even as the excitement mounts, very real challenges persist, especially as it relates to data management and data silos. Ultimately, everything comes back to data — particularly how much we can, or should, trust it. AI is only as good as the data it's trained on, and businesses need modern data architectures to truly realize the power of AI.

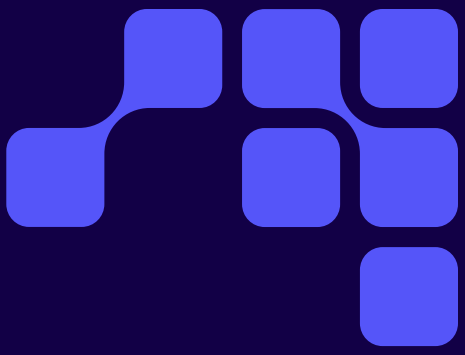
With the benefit of modern data architectures, enterprises can future-proof their AI models, drive innovation, and stand out in competitive markets.

Learn more about how Cloudera is enabling enterprises to manage and analyze data so they can unlock the full potential of Enterprise AI.

Methodology

The survey, commissioned by Cloudera and fielded by Researchscape, examines the views of 600 IT leaders based in the U.S., EMEA, and APAC regions who work at companies with more than 1,000 employees. The survey was fielded in April and May 2024. The results of this survey have been weighted to be representative of the overall GDP of surveyed countries.

About Cloudera



Cloudera is the only true hybrid platform for data, analytics, and AI. With 100× more data under management than other cloud-only vendors, Cloudera empowers global enterprises to transform data of all types, on any public or private cloud, into valuable, trusted insights. Our open data lakehouse delivers scalable and secure data management with portable cloud-native analytics, enabling customers to bring GenAI models to their data while maintaining privacy and ensuring responsible, reliable AI deployments. The world's largest brands in financial services, insurance, media, manufacturing, and government rely on Cloudera to be able to use their data to solve the impossible — today and in the future.

To learn more, visit Cloudera.com and follow us on LinkedIn and X. Cloudera and associated marks are trademarks or registered trademarks of Cloudera, Inc. All other company and product names may be trademarks of their respective owners.

CLOUDERA

Cloudera, Inc. | 5470 Great America Pkwy, Santa Clara, CA 95054 USA | cloudera.com