IMMUTA®

2024 State of Data Security Report

700+ data leaders answer questions about AI, security, visibility, and data governance.

Survey conducted by

UserEvidence

Table of Contents

Introduction & Executive Summary

Security, agility, and visibility – how are data leaders prioritizing these initiatives to protect and strengthen their businesses in 2024?

Generative AI has opened up a vast new world of productivity, possibility, and most critically, risk. AI models and use cases are being built, deployed, and restructured within a span of just months. The rate of technological innovation has left many data leaders weighing the balance between substantial gains and frightening setbacks.

In an effort to develop an authentic understanding of the current moment, we surveyed 700+ data platform and security leaders who are asking today's most pressing questions:

- How should I allocate my tech and human resources to optimize our data security?
- How can I improve my data team reporting structure?
- How are others navigating organizational, technological, and process-level security risks?
- How can my actions as a data leader directly drive better business outcomes?
- What pressing needs and data security priorities should be top of mind for me?

This survey reveals the concerns, responsibilities, and real-time decisions of data professionals who are facing an exciting yet unknown future.

Key Insights

01 AI Brings Great Excitement and Risk

In 2023, AI went from a trending topic to an organizational imperative, and data experts from across industries expect that perspective to extend far beyond 2024. As adoption continues, the savviest organizations will roll out AI data protections as part of an overall data strategy to reduce the risks of AI.

- 88% of data professionals say employees at their organizations are using AI, but only 50% say their organization's data security strategy is keeping up with the AI's rate of evolution.
- 56% of data professionals say their top concern with AI is the risk of sensitive data exposure via an AI prompt.
- Still, just 20% cite integrating AI into business processes as a top priority in the next year, compared to 35% that say implementing stronger data governance and security controls is the most significant initiative.

02 Seamless Collaboration Is a Key Component of Data Security

As more people utilize and access data across cloud data warehouses, lakehouses, and data mesh architectures, there are more stakeholders to manage. As a result, data security will continue to fall on the shoulders of everyone in the organization, not just one department or a team of IT professionals.

- 80% of respondents store at least half of their data in a cloud-based platform, and most plan to continue on this path.
- 46% of data professionals say that six to 10 people manage security at their organization, and another 33% say that 20 or more people are involved.

03 Despite Data Access Challenges, Data Leaders are Confident About Threat Response

Enabling secure data access remains a priority for data leaders, but security processes often slow down time-to-access and value. Although data leaders agree on the access management challenge, many are overconfident in their ability to sufficiently respond to threats. This signals a disconnect – if data visibility is stunted, then so is threat visibility.

- 44% of respondents say it takes a week or more to gain access to a new data set, and 56% agree that data security processes slow down access to data.
- Although access is a problem, 94% of data professionals express some degree of confidence in their ability to detect and sufficiently respond to threats. These findings may signal overconfidence in the face of access challenges.

04 Data Security Budgets Are Expected to Continue Climbing

Security spending increased in prior years and is expected to escalate. The C-Suite recognizes that data security – not just cybersecurity in general – is a top priority, and they need more resources and tools to protect their data.

- 80% of data professionals say their data protection capabilities have improved over the last year.
- 77% say their data security budgets have increased over the past year.

05 Data Governance and Security Is a Bigger Priority Than AI for 2024

Implementing stronger data governance and security controls, and modernizing data architectures, are the two top priorities – even ahead of AI. It is critical to have a data strategy before you have an AI strategy.

- Despite the buzz around AI, data governance and security was the most significant initiative respondents' organizations will take on in the next 12 months.

What significant initiatives is your organization taking on in the next 12 months?

Implementing stronger data governance and security controls	35%
Modernizing your data architecture with new concepts like data mesh	22%
Integrating AI into business processes	20%
Enabling more data self-service with data products	13%
Allocating budget towards data security training	2%
Building a bigger program to improve data privacy	2%
Other	2%

Methodology

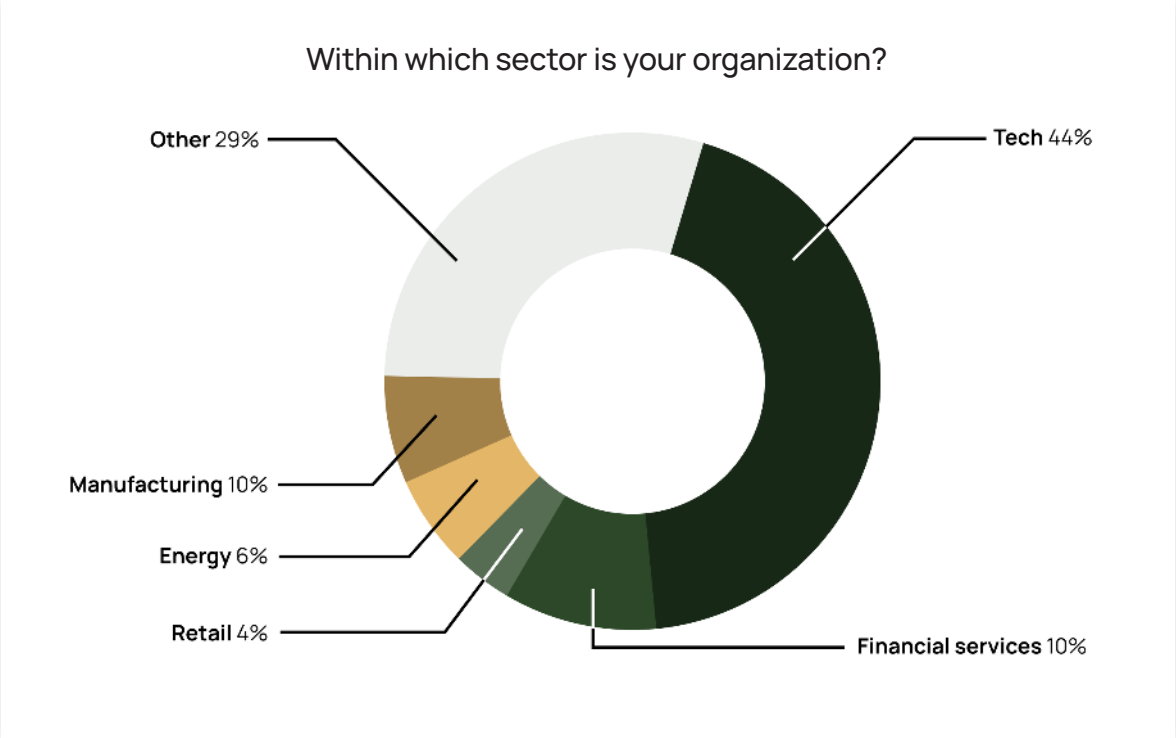
Immuta commissioned independent market research agency UserEvidence to conduct the 2024 State of Data Security Survey.

Who Are the Survey Respondents?

The study surveyed over 700 data leaders and professionals from the US, UK, Canada, and Australia.

Respondents represent global cloud-based enterprise companies across public and private sectors, including:

- Financial services
- Retail
- Manufacturing
- Technology
- Energy
- Several other sectors

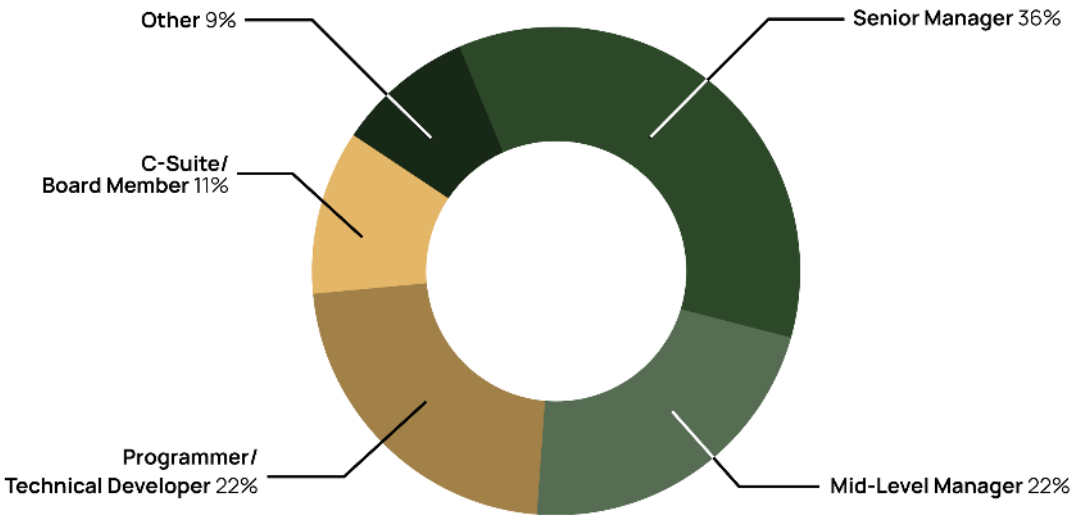


Nearly half of respondents (47%) are senior leaders or executives.

All respondents use data analytics, governance, or transformation tools in their roles, with 80% highly involved in those functions at their organization. The most common job titles included:

- IT or Engineering Director, Manager, or Head (48%)
- Data Engineer (12%)

Which of these best describes your position in the organization?



01

AI Brings Great Excitement and Risk

The rapid evolution of AI and ML has left data leaders scrambling to make high-level decisions about organizational adoption. **Nearly nine out of 10 (88%) of data professionals say that their employees are using AI, whether or not the company has officially adopted generative AI tools.** At the same time, 50% strongly or somewhat agree that their organization's data security strategy is failing to keep up with the pace of AI evolution.

While AI's presence in the workplace – and its rate of change – is exciting, data professionals are more worried about its potential risks. **Over half of respondents (56%) cite the exposure of sensitive data via an AI prompt as their greatest area of concern around AI.** Despite this, few organizations are taking steps to manage this risk through policies and governance.

> Only 21% of companies that report adopting AI have established policies for AI use in the workplace.
>
> Source: The state of AI in 2023: Generative AI's breakout year, McKinsey

Insights gathered from the survey reveal an interesting paradox: AI is simultaneously perceived as a notable security threat and a remedy.

Many data professionals are confident AI will help them become better at things like anomaly detection (44%) and phishing attack identification (46%). Virtually everyone is eager to benefit from AI, generative or otherwise, but the risks involved show that not everyone is willing to become an early adopter.

Benefits

1. Phishing Attack Identification (46%)
2. Threat Simulation & Red Teaming (44%)
3. Anomaly Detection (42%)
4. Audits & Reporting (42%)

Risks

1. Inadvertent Exposure of Sensitive Information by Employees Via Prompts (56%)
2. Unauthorized Use of Purpose-Built Models Out-of-Context (56%)
3. Inadvertent Exposure of Sensitive Information by LLMs (53%)
4. Training Data Poisoning (49%)

> "It can be difficult to ensure that sensitive information is protected from potential data leaks. Additionally, the emergence of generative AI and GPT technologies has brought about new concerns regarding data privacy and security."
>
> **Diego Souza, Chief Information Security Officer at Cummins**

Organizations must design AI-specific security strategies to match the pace of AI evolution and proliferation. Protecting data pipelines that fuel generative AI and its outputs ranks high among top priorities. Large Language Models (LLMs), which are the foundation of many generative AI tools, pose a wide range of emerging threats, such as:

- Training-data poisoning, which occurs when data fed to LLMs is manipulated
- Sensitive data disclosure by LLMs or employees
- Model theft and unauthorized use

To appropriately manage these risks, organizations need to identify, secure, and monitor the data that goes into their LLMs. Capabilities like data classification and tagging, access control, and auditing all aid in building secure and fine-tuned LLMs.

Building a vast library of AI models or use cases without considering data security could backfire.

Data leaders must implement protocols and policies that protect everyone – their organization, employees, and sensitive data. With the right policies in place, data teams can confidently position themselves to experiment and launch AI models at scale.

02

Seamless Collaboration Is a Key Component of Data Security

Data security is a responsibility that must be shouldered by everyone in the organization; it can no longer fall on one person or department. As more businesses embrace data democratization, data policies and communication guidelines will be necessary to foster safe collaboration.

The adoption of advanced frameworks, like data mesh, has accelerated democratization. By giving individual business domains responsibility for their own data governance, data mesh disperses data ownership while increasing collaboration in ways that a traditional, monolithic approach to data management never could.

It is clear, then, why data security has become a distributed function. But who is ultimately responsible for driving it?

Respondents report that **the job title most commonly accountable for data security is Data Privacy or Security Manager (19%) or Chief Technology Officer (15%)**, with numerous other roles represented as well.

As organizations grow, their structures become more complicated and they manage more data – making data security both more important and more challenging. It is not surprising, then, that data security teams are growing, too. **Nearly half (46%) of data leaders say that six to 10 people manage security for their organization, and another third (33%) say that 20 or more people are involved.**

To avoid major delays in the development cycle – or worse, insecure deployments – security is shifting left toward data teams, too. Rather than waiting to discover security issues in the production environment, organizations are pushing security testing earlier and implementing data security measures throughout development.

Shift left security forces teams and leaders to collaborate on data efforts instead of operating in a silo. When everyone aligns around data security and understands the part they play in reaching a common goal, organizations more effectively protect their data.

As essential as seamless collaboration is for data security, the industry faces a major gap when it comes to the tools professionals use.

Although security functions are shared across data security, data platforms, and legal or compliance teams, most tools are not designed with collaboration at the forefront. This makes the impact of solutions that focus on collaboration profound.

> "With Immuta, we've been able to streamline data science and engineering teamwork, dynamically adapt in real-time, and accelerate productivity."
>
> **Halim Abbas, Chief AI Officer at Cognoa**

Organizations need more than good intentions to manage security collaboratively. They also need tools and platforms that support cross-functional data security efforts and strike a good balance between accessibility and security.

03

Despite Data Access Challenges, Data Leaders Are Confident About Threat Response

With the industry-wide push to migrate data to the cloud, data leaders have found themselves between a rock and a hard place. Do they choose data security over progress and acceleration? Or do they lean into agility, becoming more vulnerable to risks and fines of noncompliance?

Organizations require tools that provide both proper data protection and the flexibility to use data to drive value.

Data Access Is Still a Considerable Security Challenge

The industry-wide challenge of data access is far from novel. The 2023 State of Data Engineering Report revealed that:

- 63% of data professionals lacked visibility into data access controls.
- Nine out of 10 data professionals said they could improve their understanding of the correlation between data access and data security.

This year's findings show that managing data access remains a major security obstacle – **33% of respondents cited a lack of visibility into data sharing and usage as their biggest security challenge.** The underlying problem is maintaining and enforcing consistent data security controls.

Recent research from Gartner supports this finding:

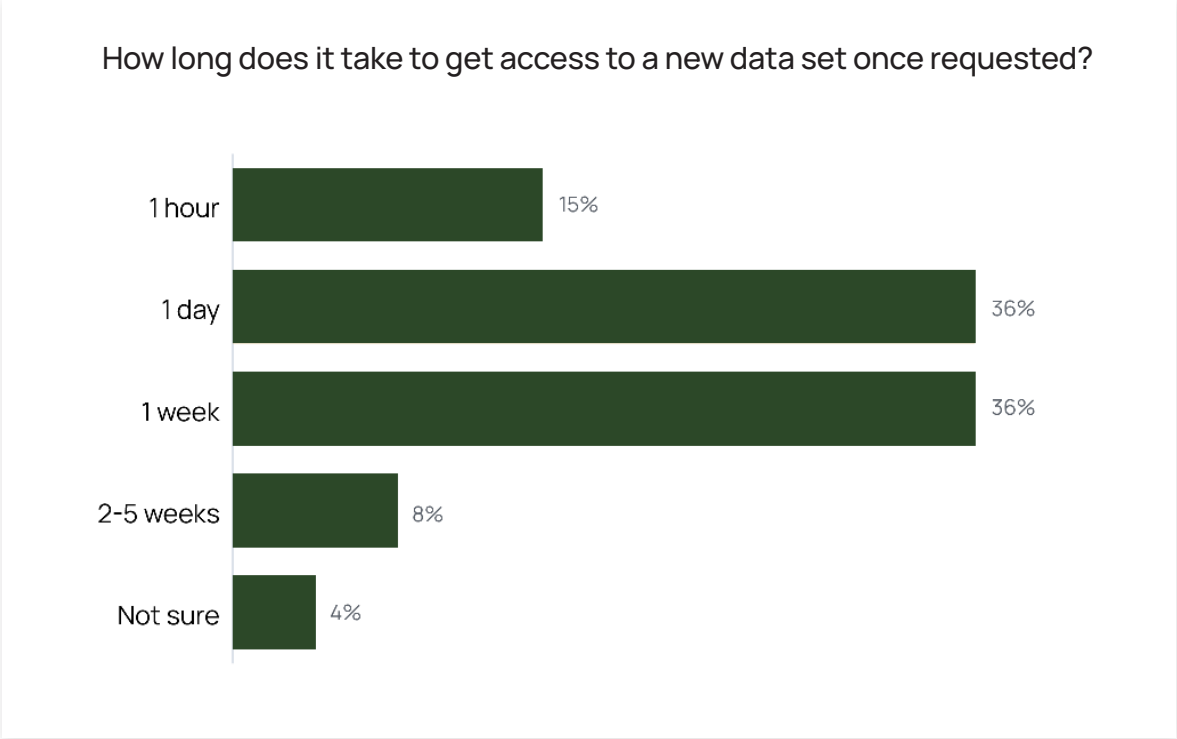
> Maintaining consistent data security is difficult because so many products provide siloed security controls, use proprietary data classification, act on specific repositories or processing steps, and do not integrate with each other. This restricts organizations' ability to identify and deploy adequate, and consistent, data security controls while balancing the business need to access data throughout its life cycle.
>
> **Gartner, Hype Cycle for Data Security, July 2023**

Yet despite these concerns, a total of 94% of data professionals feel some degree of confidence when it comes to detecting and responding to a data threat.

This reflects overconfidence from data leaders in the face of continued data access challenges. If visibility is a major issue for at least a third of data leaders, how can 94% of them proactively spot and address threats? Data professionals need streamlined security solutions in place to maintain confidence and keep data visible and protected.

> "If you don't automate your security and simplify it, the biggest problem you're going to have is some human making a mistake."
>
> **Kaj Pedersen, Chief Technology Officer at AstrumU**

This overconfidence is also directly at odds with the realities of access challenges. **Nearly half of respondents (44%) say that accessing a new data set takes a week or more.** This substantial delay inhibits organizations from using data to drive value and business outcomes.



Slowdowns and Workarounds: The Felt Impacts of Data Access Issues

Unsurprisingly, 56% of respondents agree that data security processes slow down access to data. This means that over half of organizations are sacrificing some level of data-driven value for essential security outcomes – trading agility for trust and compliance.

This leads many data users to look for workarounds – which are inherently risky, even in the most secure data environment. **More than half of respondents (57%) say that data users fall back on quick fixes to access data outside of proper data access governance practices.** This is a

problem that data leaders need to cut off at the source by improving onboarding (and offboarding) processes and building out proper access control measures.

Despite respondents' confidence in their data access management, more than half (56%) have experienced access issues that hinder business operations or outcomes.

Compliance: A Moving Target for Data Leaders

Compliance may be perceived as a peripheral challenge in comparison to data access. **Just 12% of respondents cite noncompliance as the biggest security challenge for their organization.** But the potential consequences surrounding compliance are mounting – especially in the face of unpredictable AI regulations that organizations cannot yet foresee.

Take it from the experts:

> "We are seeing a rise in compliance and privacy regulations. This trend potentially can slow the ability for an organization to use data, as well as increase the need for data access and governance solutions to prevent any potential fines or enforcement actions."
>
> **Dave DeWalt, Founder, Managing Director & CEO of NightDragon**

> "More countries are going to enforce data protection and localization regulations, slowing down the pace at which the industry is going to want to use data to drive business decisions."
>
> **Ananth Hegde, Head of Data Engineering, Commercial Bank at JPMorgan Chase**

Regulatory changes will only exacerbate existing data access challenges. Three-quarters of respondents say that they are currently subject to 10 or more regulations, which is sure to increase in the coming years.

Still, as new regulations unfold, data professionals seem optimistic about compliance. Nearly a third (32%) believe that improved data security will improve compliance for their organization.

04

Data Security Budgets Are Expected to Continue Climbing

Organizational resources and assets are flowing into the hands of data teams. 91% of respondents say they have a documented data strategy in place, and 77% say their budgets have increased. Perhaps this accounts for the 80% of data professionals reporting that their data protection capabilities are better than they were just a year ago, a trend that seems likely to continue into 2024.

Despite this, when it comes to the tools at their disposal, the majority of respondents are looking for something better.

More than half (57%) of respondents say that their tools and tech definitely need improvement.

When enterprises decide to shift data to the cloud or implement a data mesh approach, they implement security strategies to protect their organization through these changes. But by the time they're ready to go to production, they already need new security strategies in place—that's how quickly strategies need to evolve. As budgets, processes, and regulations shift over time, data teams need data security platforms and tools to shift and scale with them.

Vineeth Menon, Head of Data Lake Engineering at Swedbank, described his organization's priorities this way:

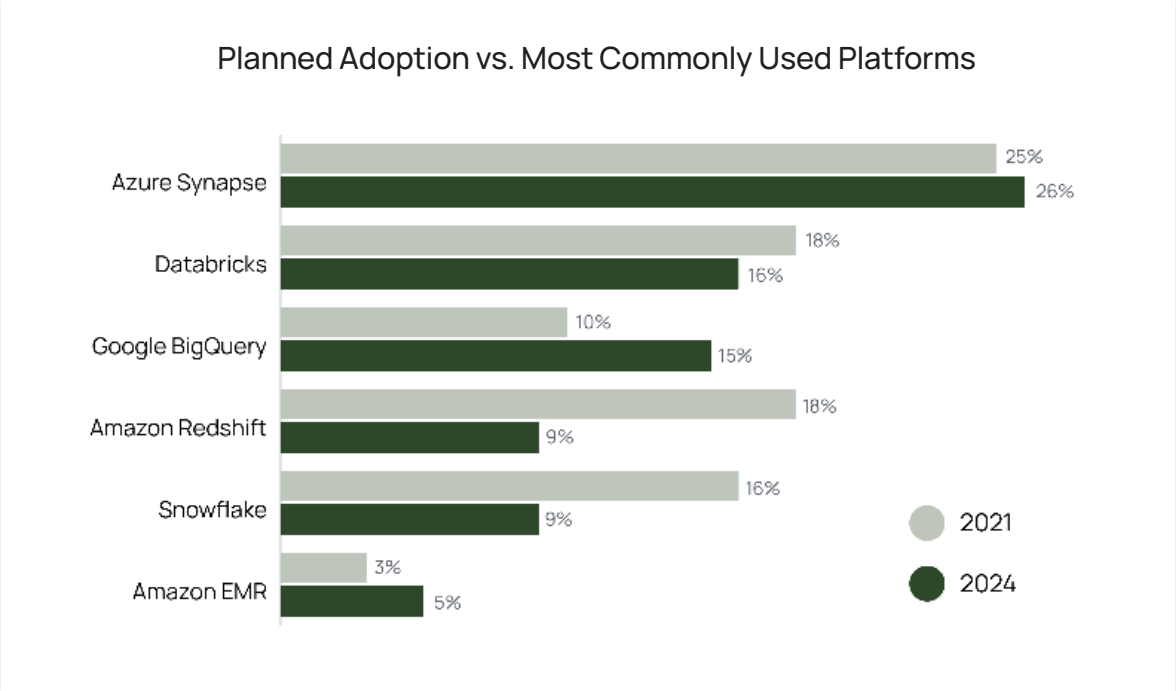
> "Swedbank needed to build an enterprise-scale advanced analytics platform that would also enforce trust in our security, management, and access to data internally while protecting our customers' assets and data."

With such multi-faceted needs for data security and access, leaders are choosing platforms that can support those needs.

Among our respondents, the most commonly used data platforms were:

1. Azure Synapse (26%)
2. Databricks (16%)
3. Google BigQuery (15%)

Comparing this data to our 2021 Impact Report, it is clear that most organizations adopted (and still use) the tools they planned on three years ago – and time to insight and collaboration remain their top priorities for cloud-based platforms.



05

Data Governance and Security Is a Bigger Priority Than AI for 2024

According to this survey, the two top priorities for next year are (1) implementing stronger data governance and security controls, and (2) modernizing data architectures. AI is number three.

The Gartner Chief Data & Analytics Officer Agenda Survey for 2023 had a similar finding, showing that the top three critical enablers for success were:

1. Governance
2. BI and reporting
3. Data literacy
4. AI/ML

Without the foundation of a strong data architecture and the proper data protections in place, it is impossible to safely integrate AI.

If polling a business audience, the responses may have been flipped given the news cycles spent on Gen AI. Gen AI is certainly exciting. However, we have to give the data professionals who responded to this survey credit for prioritizing the data hygiene required to be successful long term.

> "In the age of cloud and AI, data security and governance complexities are mounting. It's simply not possible to use legacy approaches to manage data security across hundreds of products."
>
> **Sanjeev Mohan, Principal at SanjMo**

The truth is that AI will never go from science project to production unless organizations can feel safe doing so. In another research report, Gartner predicts that, "By 2027, at least one global company will see its AI deployment banned by a regulator for noncompliance with data protection or AI governance legislation" (Gartner, Security Leader's Guide to Data Security, September 2023).

Over a third of respondents (35%) say the major initiative their organization will take on in the next year is the implementation of stronger governance and security controls.

It is perhaps worth asking whether AI itself is a driver for better security and governance. By putting tools like ChatGPT in the hands of everyday users, questions about personal data and business secrets have become hot conversations. We are entering a landscape where data and its uses are all democratized. One thing is certain – securing data across this complex landscape requires a new set of tools.

> "Organizations must foster data democratization by prioritizing providers that amplify data literacy and enable self-service data access."
>
> **The Forrester Wave™: Data Governance Solutions, Q3 2023**

2024 and Beyond: No Cons to Investing in Data Security

The pace of change has left data leaders and professionals clinging to the one truth that likely won't change for years to come: trust, security, and compliance are invaluable.

It comes as no surprise, then, that **an overwhelming majority (88%) of data leaders believe that data security will become an even higher priority in the next 12 months.** From AI advancements to shifting regulations, data professionals are rightly noting how important data security will be through the changing tides.

> "The future of our business depends on making our data and AI initiatives successful."
>
> **Jochen Kaiser, Department Lead for Data Ecosystem at Mercedes-Benz Group**

In 2024, one-third (32%) of respondents expect the top benefit of improved data security to be improved compliance. Nearly one-quarter (23%) of data professionals cite faster access to data as the top benefit.

Here are our top four recommendations for data leaders in 2024:

1. **Develop a long-term data strategy ASAP:** Frameworks and guidelines serve as the confident constant that grounds all other elements subject to change. Governance, AI, and security threats will continue to introduce unknown variables into organizational strategy. Those who invest the time and resources into high-level strategy today will have a stronger foothold in 2024.
2. **Ensure collaborative systems are in place:** Do what you can to eliminate obstacles to collaboration. It's vital that CISOs, CDOs, and other data stakeholders can work together seamlessly, no matter where they are in the world and no matter if the data is stored in cloud data warehouses, lakehouses, or data mesh architectures. The ROI on business intelligence and data-informed decisions will only increase.
3. **Craft a plan to assess and address potential data access issues:** Data security is dynamic. It requires consistent monitoring, maintenance, and oversight. Make it a priority in 2024 to speed up time-to-access for your overall organization without compromising governance. This is a tall order, but teams who strike a balance between data visibility, risk mitigation, and accessibility will not only be the most secure, but also the most efficient.
4. **Place the right big bets on your data platform:** The savviest data professionals are not trying to juggle security, agility, and visibility on their own. Simplify and improve your data security with an industry expert like Immuta. Our platform enables organizations to unlock value from their cloud data by protecting it and providing secure access.

Think Immuta might be right for your organization? Spend 29 minutes with us.

Glossary

Data access governance / Data access management: Data access governance and data access management are often used interchangeably. Both terms refer to an organization's policies and procedures that facilitate data access. There are particular models for data access, such as Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC). This term refers to the implementation of those models and the management of the policies, including how they are stored, executed, and monitored.

Data mesh: Data mesh is a modern data architecture design that decentralizes data ownership, leaving data governance and quality in the hands of self-contained domains. Built to meet the needs of hybrid and multi-cloud data environments, data mesh reduces data silos and bottlenecks, and puts data in the hands of those who need it.

Generative AI: Generative AI is a form of artificial intelligence that uses generative models to create and output various content types, including text, images, graphics, audio, and video.

Large Language Model: A Large Language Model (LLM) is a type of machine learning model known for its natural language processing (NLP) capabilities. LLMs are trained using massive data sets, and can both understand humans and generate conversational responses to questions or prompts.

Machine Learning: Machine learning (ML) is a subset of artificial intelligence (AI) that uses data and algorithms to train machines to learn as humans do. ML plays an emerging role in the data science field, as algorithms learn to review large data sets then identify patterns, predict outcomes, and make classifications.

Shift left: The shift left approach to security introduces security measures and testing early in the development cycle to avoid major delays or expensive errors due to security or compliance gaps.

About Immuta

Immuta enables organizations to unlock value from their cloud data by protecting it and providing secure access. The Immuta Data Security Platform provides sensitive data discovery, security and access control, data activity monitoring, and has deep integrations with the leading cloud data platforms. Immuta is now trusted by Fortune 500 companies and government agencies around the world to secure their data. Founded in 2015, Immuta is headquartered in Boston, MA.

IMMUTA®

25 Thomson Place, 4th Floor, Boston, MA 02210 immuta.com (800) 655-0982