privacera

E-BOOK

AWS

Mastering Data Access

Why DIY Cloud Governance Falls Short

contents

INTRODUCTION

It’s time for CIOs and CISOs to confront the data security and governance complexities of their cloud environments.

Organizations traditionally managed singular, relatively straightforward systems. Today, with hybrid and multi-cloud infrastructures, they face a tangled web of diverse storage, compute, and consumption technologies. The challenge of securing and governing data across this landscape is immense.

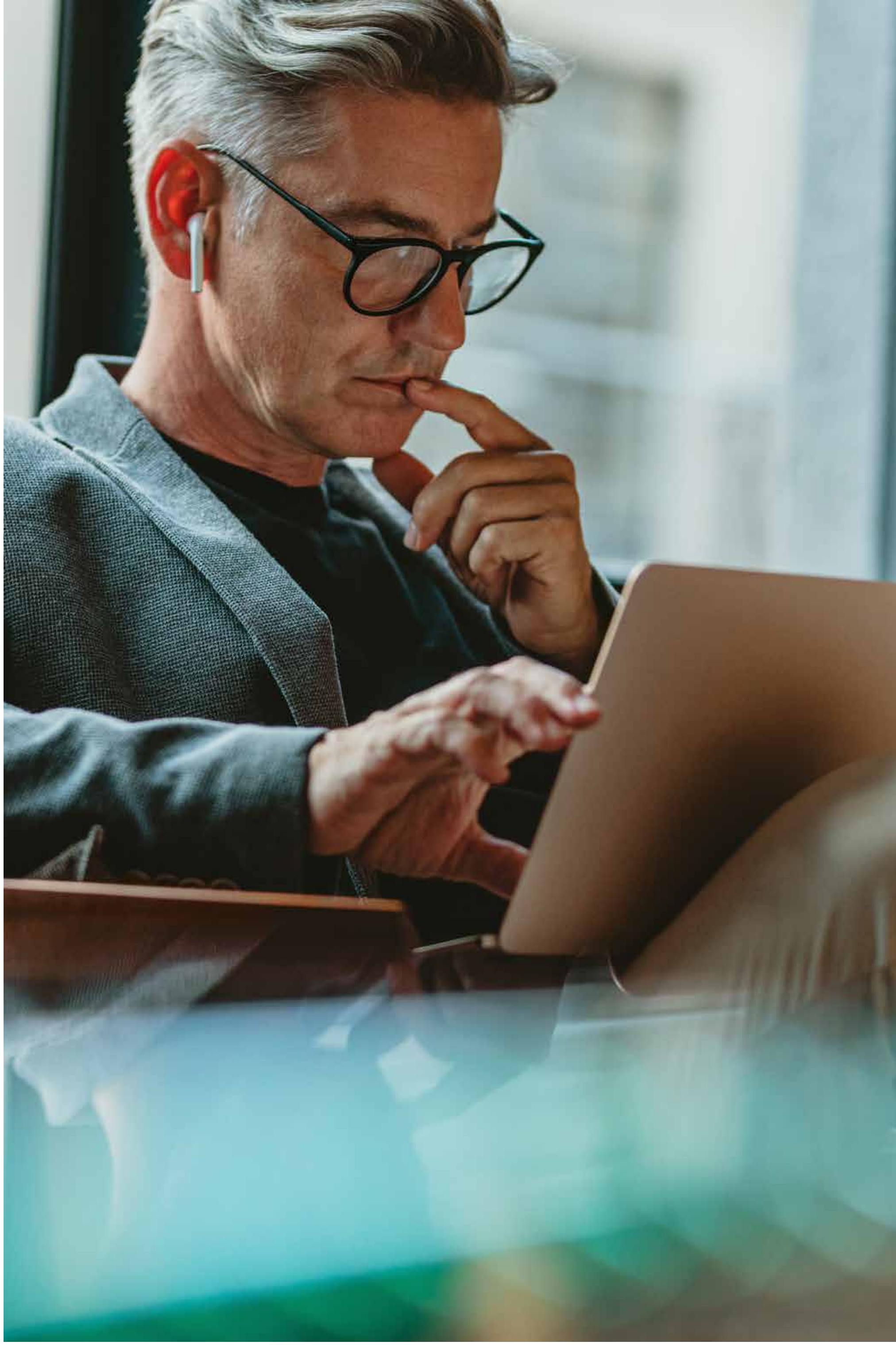
Unfortunately, organizations have resorted to solving the security and governance complexity leveraging the native options provided by AWS, who often claim to solve data governance problems pertaining to not only the single cloud, but also hybrid, multi-cloud environments. However, organizations quickly realize that this Do-It-Yourself (DIY) solution creates new data security and access management challenges within the single cloud. Then for the hybrid, multi-cloud environment, DIY solutions only work within the specific cloud vendor ecosystem, and they have the burden of integrating different data sources and managing security and governance across their diverse environment. As organizations dig deeper, the stark reality emerges: governing data across hybrid, multi-cloud is far more complicated. Furthermore, these solutions are not as scalable and flexible as one would expect.

Most organizations who had resorted to DIY solutions with a single cloud vendor discovered that their governance and compliance needs are unmet. For this reason, CIOs and CISOs must take action to streamline the security and governance challenges posed by any (single, hybrid and/ or multi) cloud environment and bolster protection before it’s too late. They need a unified data security platform.

PART 1

Challenges for AWS-Only Environments

Amazon Web Services (AWS) is a leading cloud platform for organizations migrating data and processing workloads to the cloud. One of AWS's main advantages is its flexibility and the wide range of storage and processing options available to suit different workloads. Central to AWS's storage offerings is Amazon S3, an object store capable of storing various file types. A key feature of S3 is the concept of "buckets," which are buckets that organize these objects. Organizations face numerous security and governance challenges with AWS-only environments.

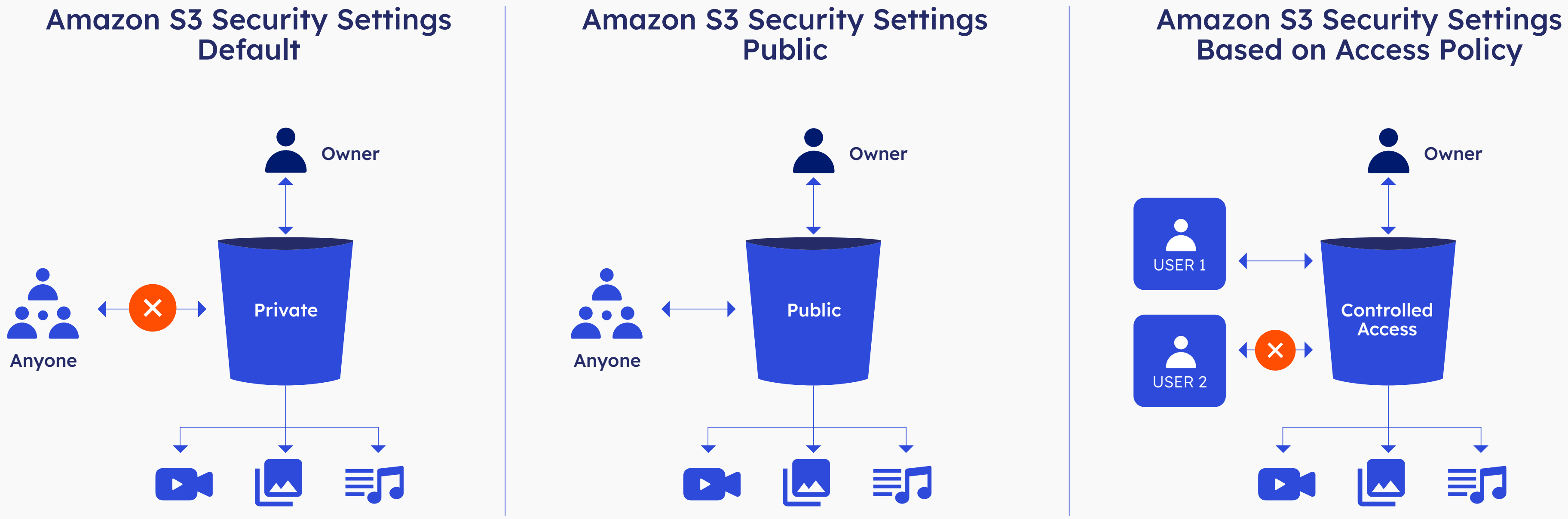


CHALLENGE #1

Over-permissioning due to IAM role-based access impose a security risk

AWS' security controls are controlled by IAM Roles. Essentially, a user gets assigned a role that then determines authorization levels.

A major issue arises with AWS S3 IAM-based access control that operates at the bucket level, meaning your users with access to a bucket can see all objects within it. If you have a bucket that contains 1,000 objects, every user with access to that bucket will typically gain visibility into all 1,000 objects - even those that should remain restricted. This results in over-permissioning risks, where users may access sensitive files that aren't relevant to their role. Managing this at scale across thousands of S3 buckets and objects becomes a significant operational challenge.

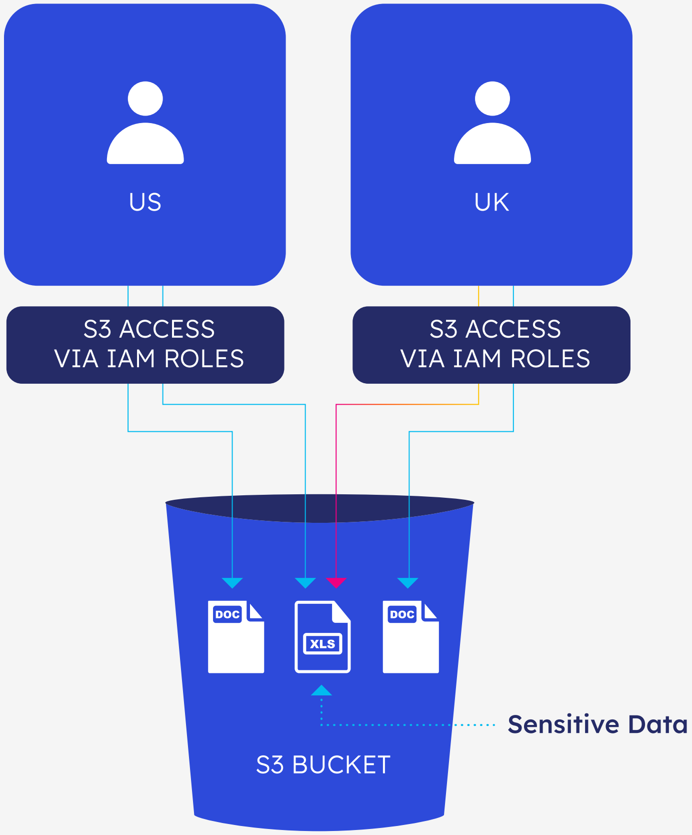

Amazon S3 Security Settings Default
Owner
Anyone
Private
Amazon S3 Security Settings Public
Owner
Anyone
Public
Amazon S3 Security Settings Based on Access Policy
Owner
USER 1
USER 2
Controlled Access

A major issue arises with AWS S3 IAM-based access control that operates at the bucket level, meaning your users with access to a bucket can see all objects within it.

UNWANTED SCENARIO EXAMPLE #1:

Absence of Granular Access Controls Risks HR Data Exposure

The US team has data for all the employees and it also includes sensitive data. All this data is in a bucket. The same bucket is shared with the UK team. The UK team also puts its data and uses it. Unfortunately, there is no control over the access to specific data (files, data, column etc), and the UK team ends up having visibility into US data – including the sensitive data.



Per Compliance mandate, the UK shouldn't have visibility into US data.

Here the UK team HAS visibility into sensitive US data.

✕ **GOVERNANCE**

✕ **COMPLIANCE**

✕ **GDPR, CCPA**

SOLUTION #1

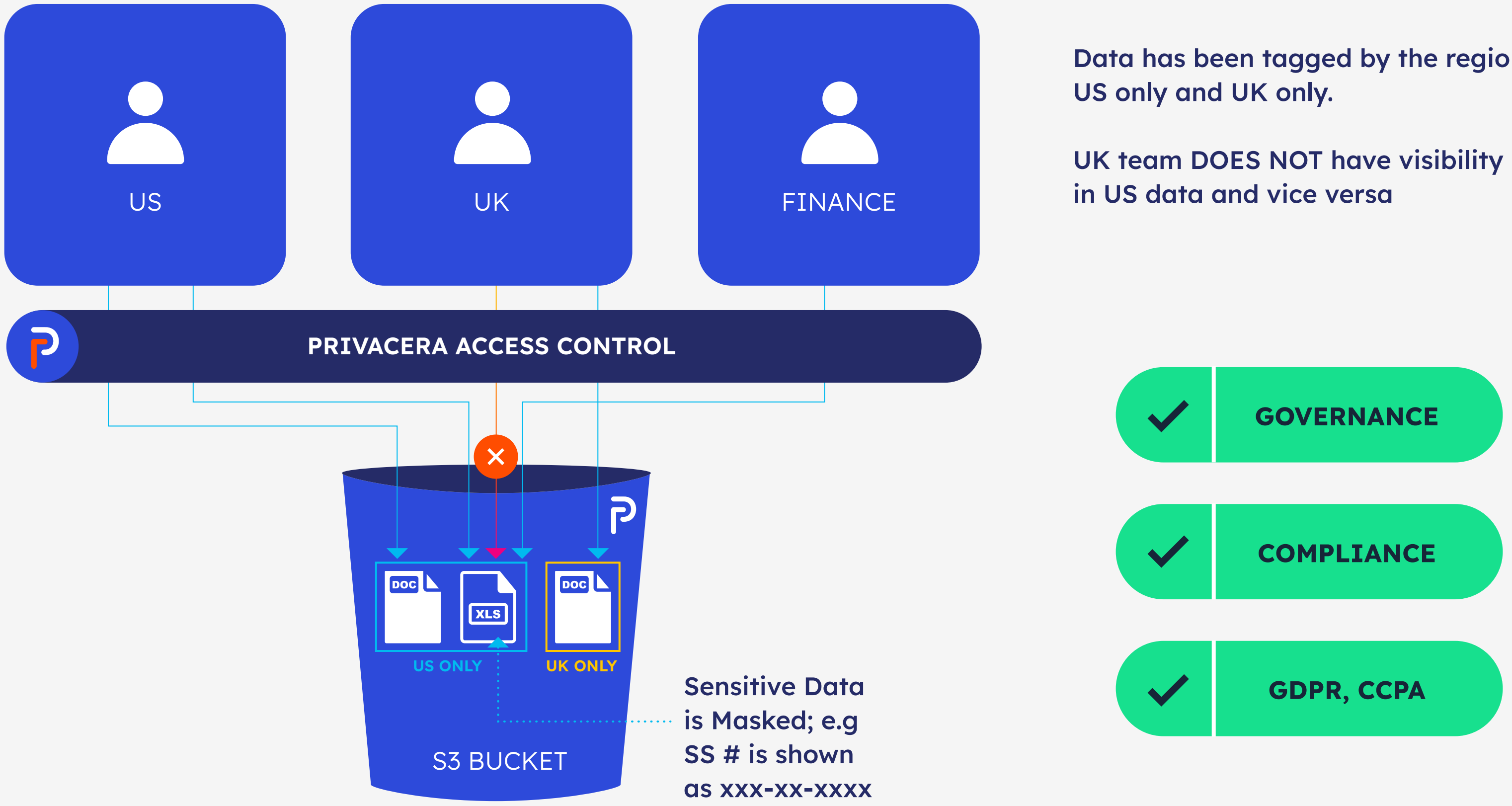
Privacera centralizes access to S3 bucket with ABAC and object-level controls to provide fine-grained access

With Privacera, you can centralize access management for S3. This helps you eliminate the need for thousands of IAM roles. You can classify or tag data at bucket, folder or object level and policies can be created to provide access based on those classification. After which, attribute-based access control (ABAC) model can allow you to dynamically set permissions based on user attribute, data attribute, group affiliations, and roles - giving you fine-grained access control. This fine-grained access control ensures only authorized users can access specific objects, reducing over-permissioning risks.

IDEAL SCENARIO EXAMPLE #1:

Privacera Safeguards Sensitive Data with Attribute-Based Access Control

The US team has sales and marketing data - including sensitive data in the same bucket as the Sales team. Though the bucket is shared between the two members from the two countries, with Privacera the UK team only has access to the data that is granted to them. The data steward tagged the US data as "US only" data and UK data as "UK only" data, and created a policy to provide access to people within the locations. Then through attribute-based access control (ABAC) model, Privacera controls access based on users location and, hence, the sensitive data is only visible to the US team and not the UK team. Furthermore, Privacera also enabled the marketing team to share some datasets that have sensitive data with the finance team by masking (e.g. xxx-xx-xxxx) it.



Data has been tagged by the region:
US only and UK only.

UK team DOES NOT have visibility
in US data and vice versa

Sensitive Data
is Masked; e.g
SS # is shown
as xxx-xx-xxxx

CHALLENGE #2

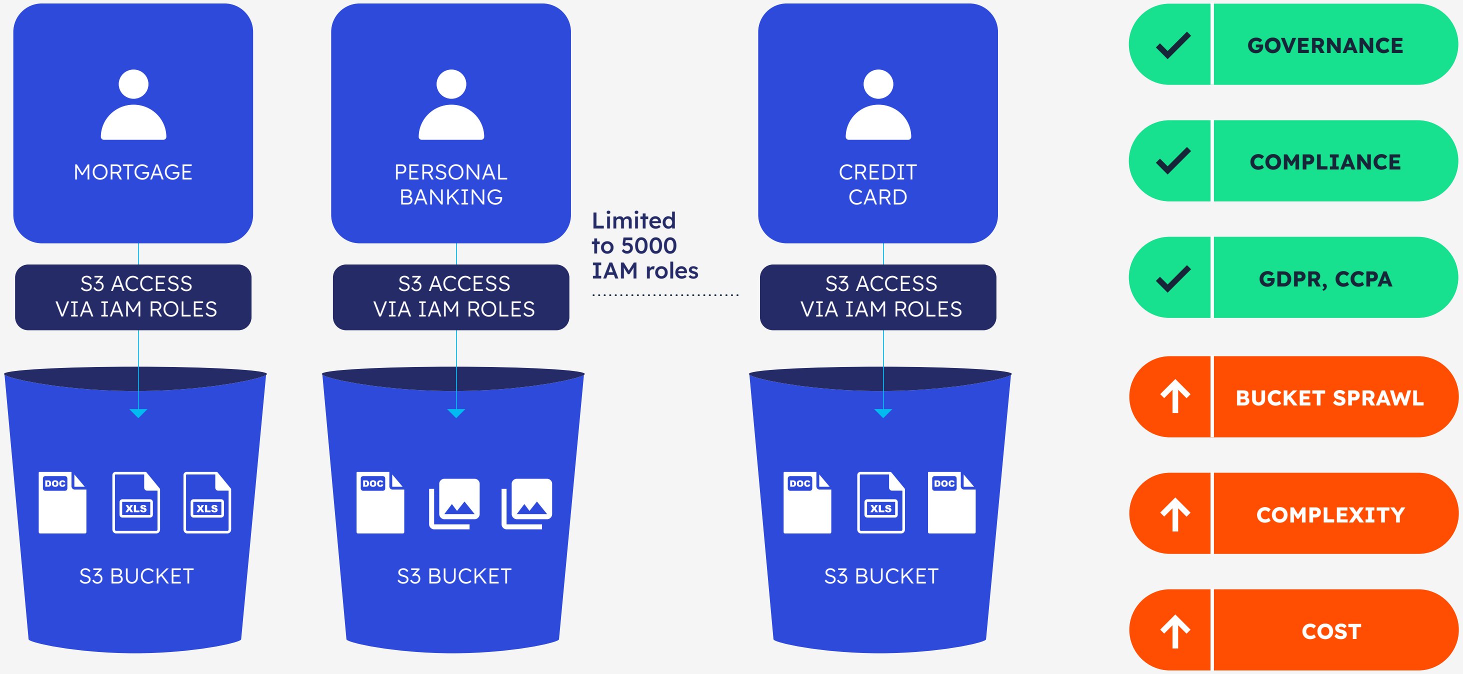
Creating separate buckets for sensitive data results in role sprawl and complexity

To limit access to sensitive objects, you have to create separate buckets to isolate sensitive files. In addition to the new bucket, with this approach you have to also create and manage additional IAM roles. Over time, this leads to bucket sprawl, role sprawl, and a complex system that's difficult to maintain. The complexity increases your (and your team's) administrative overhead, as each new role or bucket requires manual intervention to configure access and permissions, creating unnecessary operational burden. In addition, there is a hard limit to number characters (2000) you can have in a policy and number of IAMs roles (5000) in a domain.

UNWANTED SCENARIO EXAMPLE #2:

Financial institute ends up creating multiple buckets

In a financial institute, Bank of Narnia, they have different departments, such as mortgage, personal banking, credit card etc. Due to compliance, such as GDPR and CCPA, the institute has to ensure that all this data is separate from each other. So, the IT team started creating buckets for each department, and with each bucket created a separate IAM role. However, this sprawl of buckets and roles created an unnecessary burden on them. As a result, they ended up managing 1000s of buckets and roles, which has become a daunting task. The IAM roles increased to 7000 due to different projects and they ended hitting the limit of 5000 within the domain.



SOLUTION #2

Privacera eliminates bucket sprawl with object-level access control and automated permissions

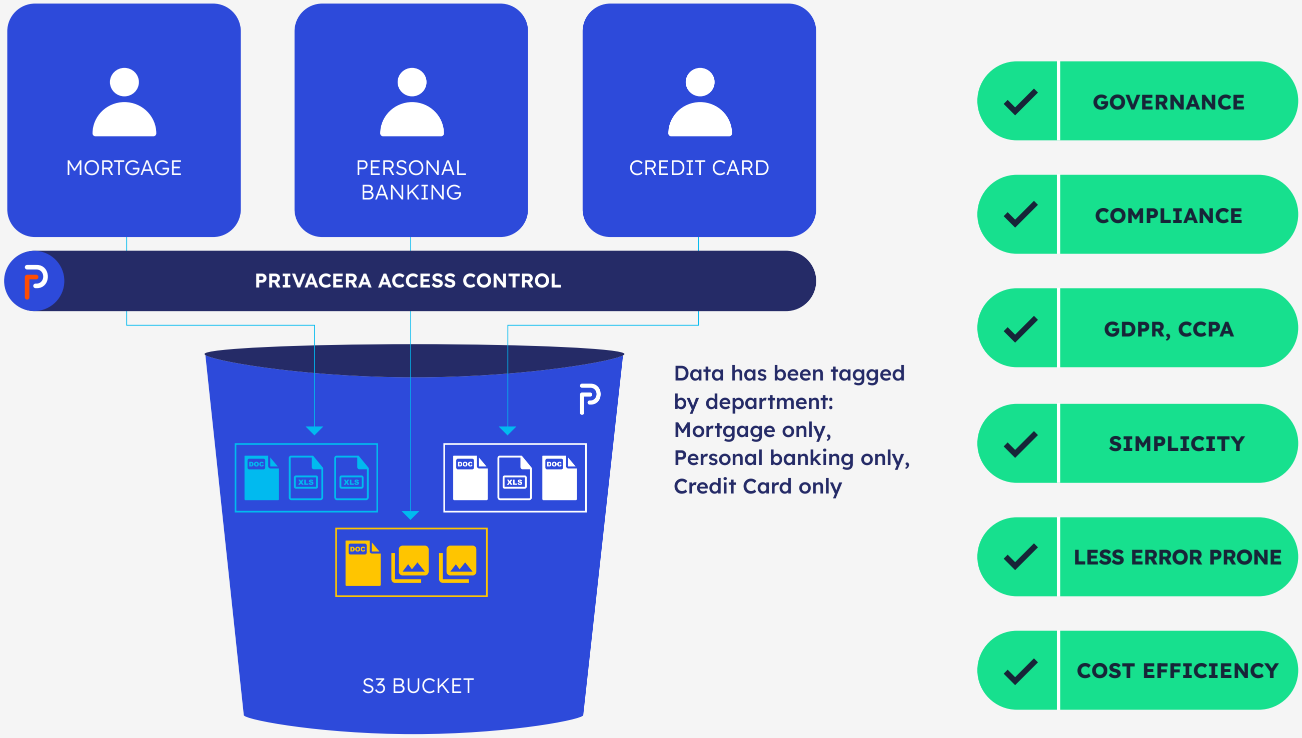
With Privacera, you can eliminate the need for separate buckets to isolate sensitive data. This is because Privacera provides you a platform where you can get centralized dynamically controlled access at object-level. Instead of creating new IAM roles and buckets for each use case. Data stewards can classify or tag data at bucket, folder or object level and policies can be created to provide access based on those classification. After which, attribute-based access control (ABAC) model can allow you to dynamically set permissions based on user attribute, data attribute, group affiliations, and roles. Privacera then helps you manage permissions through its automated access model - streamlining access control and governance to specific groups or users. This approach helps you to significantly reduce bucket and role sprawl and never hit the character and/or bucket limit.

IDEAL SCENARIO EXAMPLE #2:

Financial institute controls the bucket sprawl

The Bank of Narnia controlled bucket and role sprawl by putting all the data from all the departments in the same bucket. They leveraged Privacera to manage data access by individuals (user attribute). Data stewards tagged the Mortgage data as “Mortgage only” data, Personal banking data as “Personal banking only” data, and Credit card data as “Credit card only” data, and created a policy that users within respective departments can only access data assigned to them. Then through attribute-based access control (ABAC) model, Privacera controls access based on the user’s attributes.

Hence, only the assigned data is visible to each department member. That is, though the same bucket is shared among different departments, each member of the mortgage department only has access to the data granted to the mortgage department, each member of the personal banking department only has access to the data granted to the banking department, and so on and so forth. This reduction of the number of buckets and roles helped them overcome the hard limits of maximum buckets and roles they could possess.



Data has been tagged by department: Mortgage only, Personal banking only, Credit Card only

- ✓ GOVERNANCE
- ✓ COMPLIANCE
- ✓ GDPR, CCPA
- ✓ SIMPLICITY
- ✓ LESS ERROR PRONE
- ✓ COST EFFICIENCY

CHALLENGE #3

Manually granting access control hampers scalability and efficiency

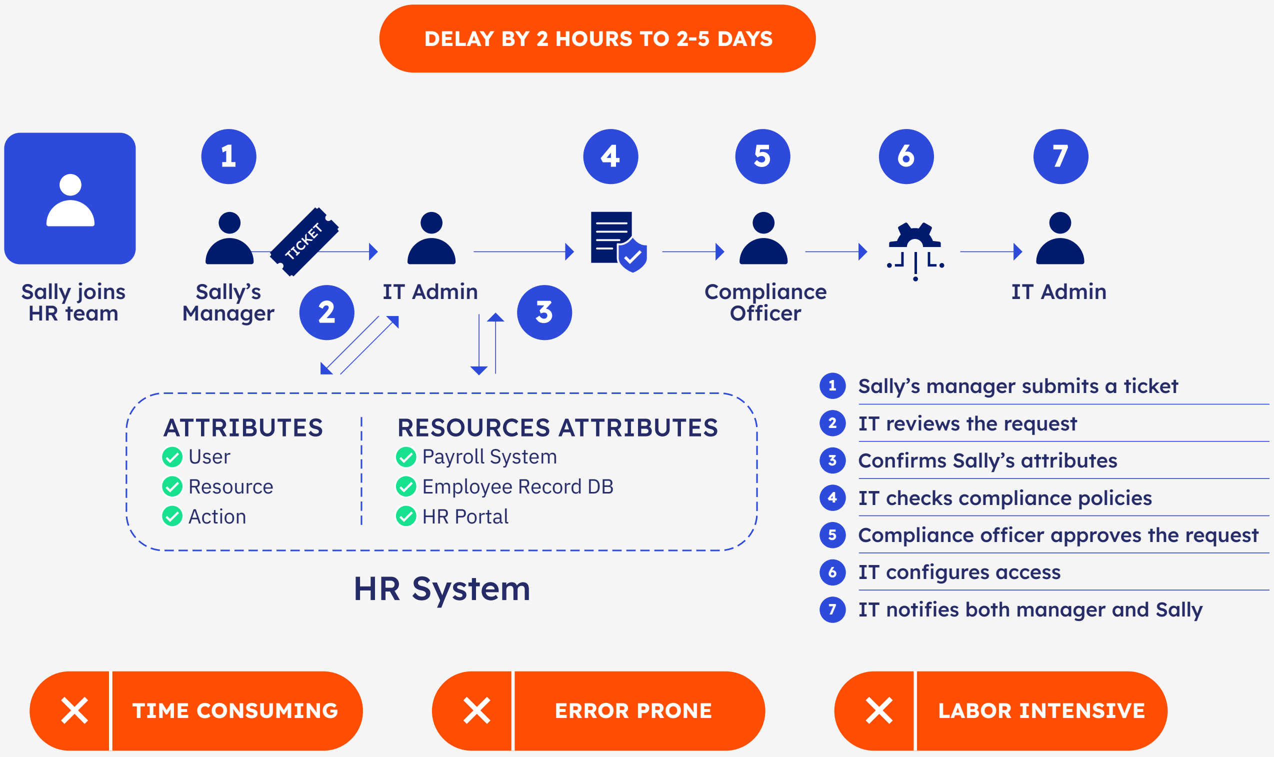
Organizations managing large-scale data environments face significant challenges with manual processes for access control and governance. The reliance on ticketing systems for access requests in platforms like S3 creates bottlenecks, delays, and a heavy administrative burden. Each new request requires manual review and approval, leading to slower data access, reduced productivity, and a higher likelihood of errors. In some cases companies could take 2 hours to 2-5 days.

Additionally, integrating new objects, datasets, and tables into governance frameworks compounds the problem. Every new asset requires manually creating roles, permissions, and policies, which increases operational complexity. This labor-intensive approach results in delays, misconfigurations, and potential security vulnerabilities, making it difficult to scale efficiently while maintaining compliance and security standards.

UNWANTED SCENARIO EXAMPLE #3:

The Hassle of Granting Data Access: A New Employee's Onboarding Challenge

Sally, a new member of the HR team, joins the organization, and her manager submits a ticket via the IT department's ticketing system requesting access to HR-related data and tools. The IT team reviews the request, confirms Sally's role, and seeks additional clarification on specific data and systems she requires access to. The manager responds, detailing that Sally needs access to the payroll system, employee records database, and the HR portal for managing benefits. The IT team checks compliance policies and ensures the requested access aligns with her role. They may reach out to the HR compliance officer for approval if sensitive data is involved. After final approval, the IT team configures access and notifies both the manager and Sally. Sally then tests her access, and if any discrepancies arise, she or her manager may need to follow up for adjustments, closing the ticket once all issues are resolved.



DELAY BY 2 HOURS TO 2-5 DAYS

Sally joins HR team

1 Sally's Manager

TICKET

2 IT Admin

3

4

5 Compliance Officer

6

7 IT Admin

ATTRIBUTES
- ✓ User
- ✓ Resource
- ✓ Action

RESOURCES ATTRIBUTES
- ✓ Payroll System
- ✓ Employee Record DB
- ✓ HR Portal

HR System

1. Sally's manager submits a ticket
2. IT reviews the request
3. Confirms Sally's attributes
4. IT checks compliance policies
5. Compliance officer approves the request
6. IT configures access
7. IT notifies both manager and Sally

✕ TIME CONSUMING

✕ ERROR PRONE

✕ LABOR INTENSIVE

SOLUTION #3

Privacera increases efficiency by automating data access, governance, and compliance

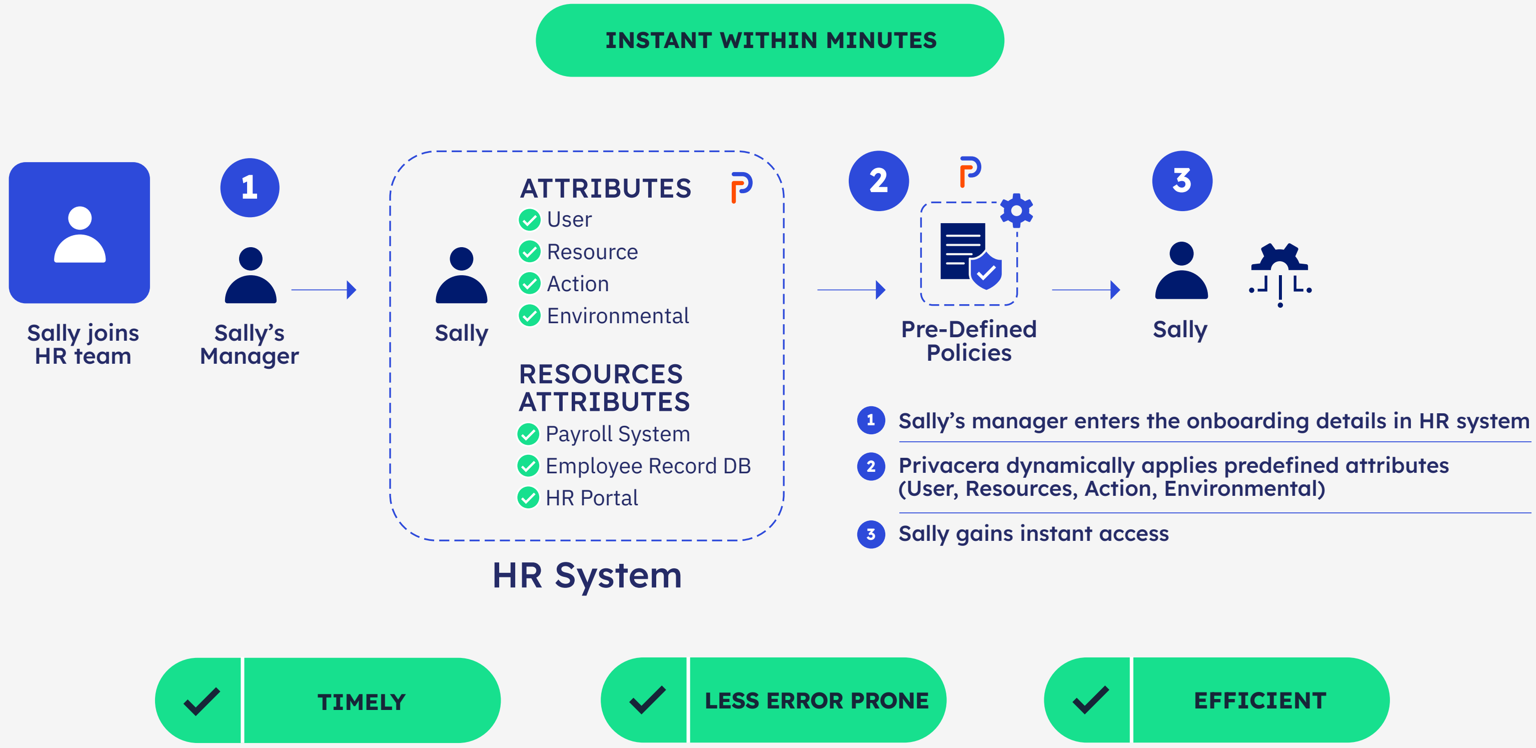
Privacera streamlines access control and governance by automating the management of permissions and policies for datasets in S3. Its integrated approach eliminates the need for manual ticketing systems and manual policy creation, reducing overhead and accelerating data access. By dynamically assigning permissions based on attributes (User, Object, Action, Environmental), Privacera ensures that access is granted efficiently and securely. Furthermore, new objects, datasets, and tables are automatically governed as soon as they are created, with existing rules applied seamlessly. This automation eliminates delays in onboarding new people or adding new assets, and ensures consistent governance without sacrificing security or compliance.

With Privacera, organizations can manage the exponential growth of data while maintaining operational efficiency. Teams can focus on innovation and strategic tasks rather than administrative work, enabling scalable and secure data environments that adapt effortlessly to evolving business needs.

IDEAL SCENARIO EXAMPLE #3:

Streamlined Onboarding: How Privacera Simplifies Secure Data Access for New Employees

When Sally joins the HR team, the ideal scenario using Privacera streamlines and automates her access to necessary datasets and systems without relying on manual ticketing processes. Once Sally's onboarding details, such as her role, group memberships, and attributes, are entered into the HR system, Privacera dynamically applies predefined access policies aligned with the HR role. These policies grant her secure access to specific datasets in S3, such as payroll and employee records, while ensuring compliance with governance rules. As Privacera automatically governs new datasets and applies existing rules to them, Sally gains instant, secure access to any new HR datasets created after her onboarding. This approach eliminates delays, reduces administrative overhead, and ensures consistent governance, enabling Sally to start contributing to her team immediately while maintaining robust security and compliance standards.



1. Sally's manager enters the onboarding details in HR system
2. Privacera dynamically applies predefined attributes (User, Resources, Action, Environmental)
3. Sally gains instant access

CHALLENGE #4

Inefficiencies in Manual Access Control for Adding New Datasets

Adding new datasets to large-scale data environments creates challenges due to manual, labor-intensive processes. Each dataset requires defining roles, permissions, and policies, involving coordination between IT administrators, compliance officers, and data stewards. Reliance on ticketing systems exacerbates delays, as requests undergo lengthy reviews and approvals, often taking hours to several days. These inefficiencies slow access to critical data and reduce productivity across teams.

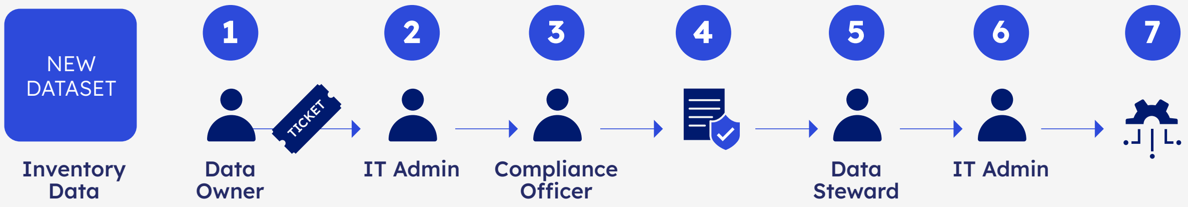
The manual nature of these workflows also increases the risk of errors, misconfigurations, and security vulnerabilities, especially with sensitive information. Scaling these processes efficiently while maintaining compliance becomes a significant challenge.

UNWANTED SCENARIO EXAMPLE #4:

Manually Addition of Dataset at a Large Retail Compromises Security & Compliance and Creates Inefficiencies

In a large retail company, Tarmart, adding a new dataset to its system involves multiple stakeholders, including compliance officers, IT administrators, and data stewards, each playing a critical role. The process begins with the data owner submitting a request, detailing sensitive information like PII or email data within the dataset. The IT administrator reviews the request and forwards it to the compliance officer, who determines the appropriate access policies based on the type of data and the user groups needing access. Once the compliance officer establishes the policy, they pass it to the data steward, who ensures the dataset is managed according to organizational standards. Finally, IT administrators upload the dataset and grant access based on the defined policies, acting as facilitators. Though this collaborative workflow ensures security and compliance, it can also introduce inefficiencies, especially when managing multiple stakeholders and manual approvals.

DELAY BY 2 HOURS TO 2-5 DAYS



NEW DATASET

Inventory Data

1 Data Owner

TICKET

2 IT Admin

3 Compliance Officer

4

5 Data Steward

6 IT Admin

7

1. Data Owner submits a ticket to assign policy for the new dataset – including information of sensitive data (PII, Email) in the datasets
2. IT Administrator takes the request and submits it to the compliance officer
3. Compliance Officer understands the class of people to can access the data
4. Creates a new policy
5. Sents the update to the Data Steward
6. Data Steward enforces policies and manages (data integrity, governance, monitor) the data
7. Admin finally uploads the data and grants assess to data (facilitator)

✕ TIME CONSUMING

✕ ERROR PRONE

✕ LABOR INTENSIVE

SOLUTION #4

Automating Data Access, Governance, and Compliance at Scale

Privacera simplifies the integration of new datasets into S3 environments by automating permissions and policy management. Unlike manual ticketing systems and policy creation, Privacera dynamically assigns permissions based on attributes such as User, Object, Action, and Environmental factors.This ensures efficient, secure access without manual intervention. New datasets, objects, and tables are automatically governed upon creation, seamlessly inheriting existing rules. This eliminates delays typically associated with adding new assets while maintaining consistent governance, security, and compliance standards.

By automating dataset governance, Privacera enables organizations to manage data growth without the administrative burden. Teams can allocate more time to strategic initiatives rather than repetitive tasks, ensuring scalable and secure data environments that adapt to evolving business demands with ease.

IDEAL SCENARIO EXAMPLE #4:

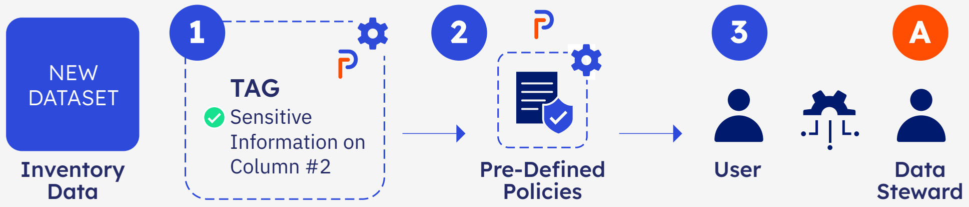
Privacera Ensures Access, Security, and Compliance at Scale by Automating Access to New Datasets at Tarmart

At Tarmart when the data owner added a new dataset, e.g to Inventory Data, privacera automatically detected that a new dataset was added. Next, it identified sensitive information based on data and metadata (e.g. schema, format). Based on this information, Privacera then applies the necessary tags, and makes it available to the system by making it public. Existing policies are leveraged to apply permissions to the data. In this case, Privacera ensured that the newly created "dataset" in inventory Data in S3 was instantly governed with existing policies.

As a result, the following policy was enforced: "Mask sensitive information for non-HR and non-US members". Instead of relying on ticketing systems and manual approvals, Privacera dynamically applies predefined access controls based on user roles, group memberships, and compliance requirements. This guarantees that authorized users can securely access new Inventory Data without disrupting workflows.

This automation reduces administrative overhead at Tarmart and enhances operational efficiency - which enables teams to make immediate use of new Inventory Data for business insights and decision-making.

INSTANT WITHIN SECONDS



1. Privacera detects a new dataset, identifies sensitive information based on data and metadata (e.g schema, format), assigns tag and makes data available to system
2. Privacera leverages pre-defined polices to apply permission: Mask sensitive information for non-HR and non-US members
3. Users are granted "birth right access" to the dataset

A. Data Steward manages (data integrity, governance, monitor) the data

✓ TIMELY

✓ LESS ERROR PRONE

✓ EFFICIENT

CHALLENGE #5

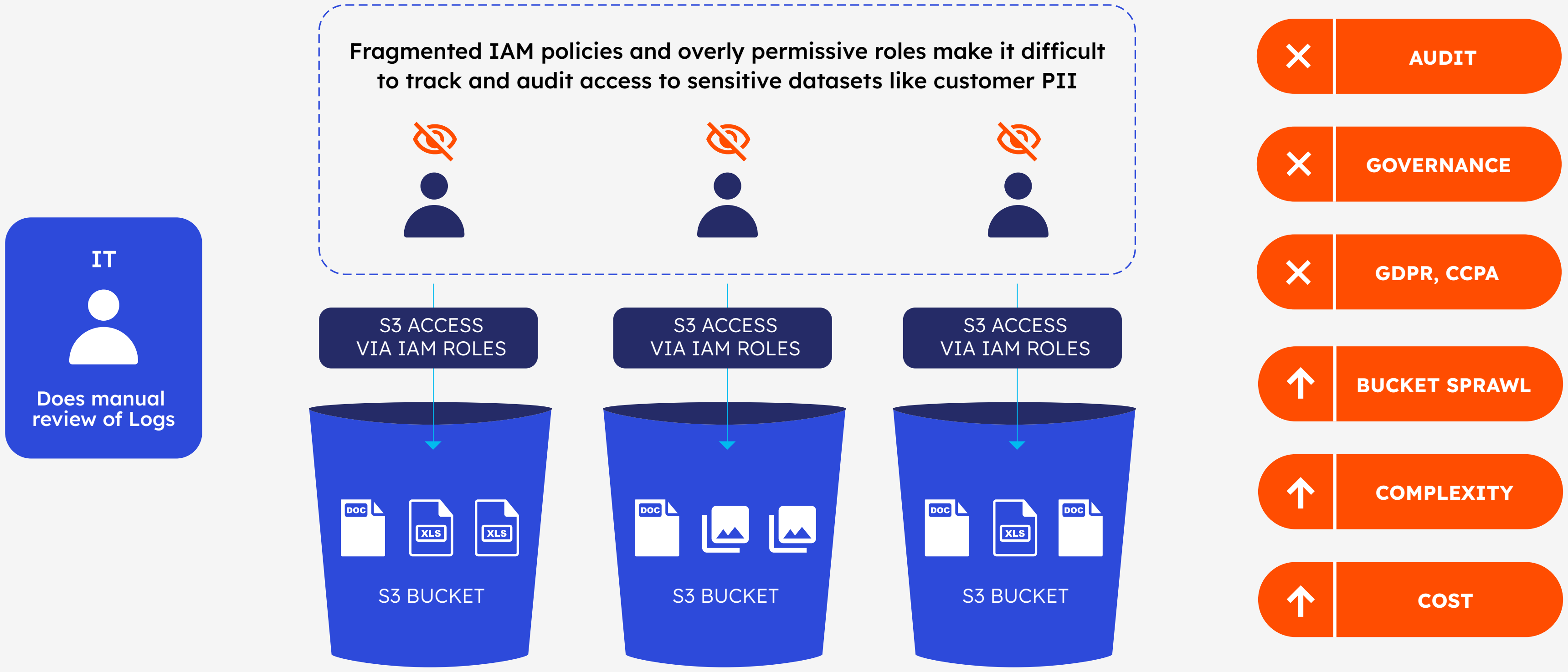
Lack of visibility and auditability of access and permissions

With AWS S3's default IAM approach, it is difficult for you to track and audit who has access to which datasets. IAM permissions can become overly complex, and visibility into which users have access to what data is limited. Without clear visibility into access permissions and activity, your organization will struggle to enforce governance and compliance mandates like GDPR, HIPAA, and CCPA. Audit preparation becomes resource-intensive and prone to human error.

UNWANTED SCENARIO EXAMPLE #5:

Global Retail's Data Governance Struggle: Tackling Compliance and Access Challenges in AWS S3

A global retailer, Tarmark, faces challenges managing data access across its AWS S3 infrastructure. To ensure governance, the IT team created 100s of buckets, which resulted in 100s of IAM roles. The result is that fragmented IAM policies and overly permissive roles have made it difficult to track and audit access to sensitive datasets like customer PII, complicating compliance with regulations like GDPR. During audits, the IT team does manual reviews of logs – resulting in them not only taking time to resolve the compliance, but also having incomplete information.



SOLUTION #5

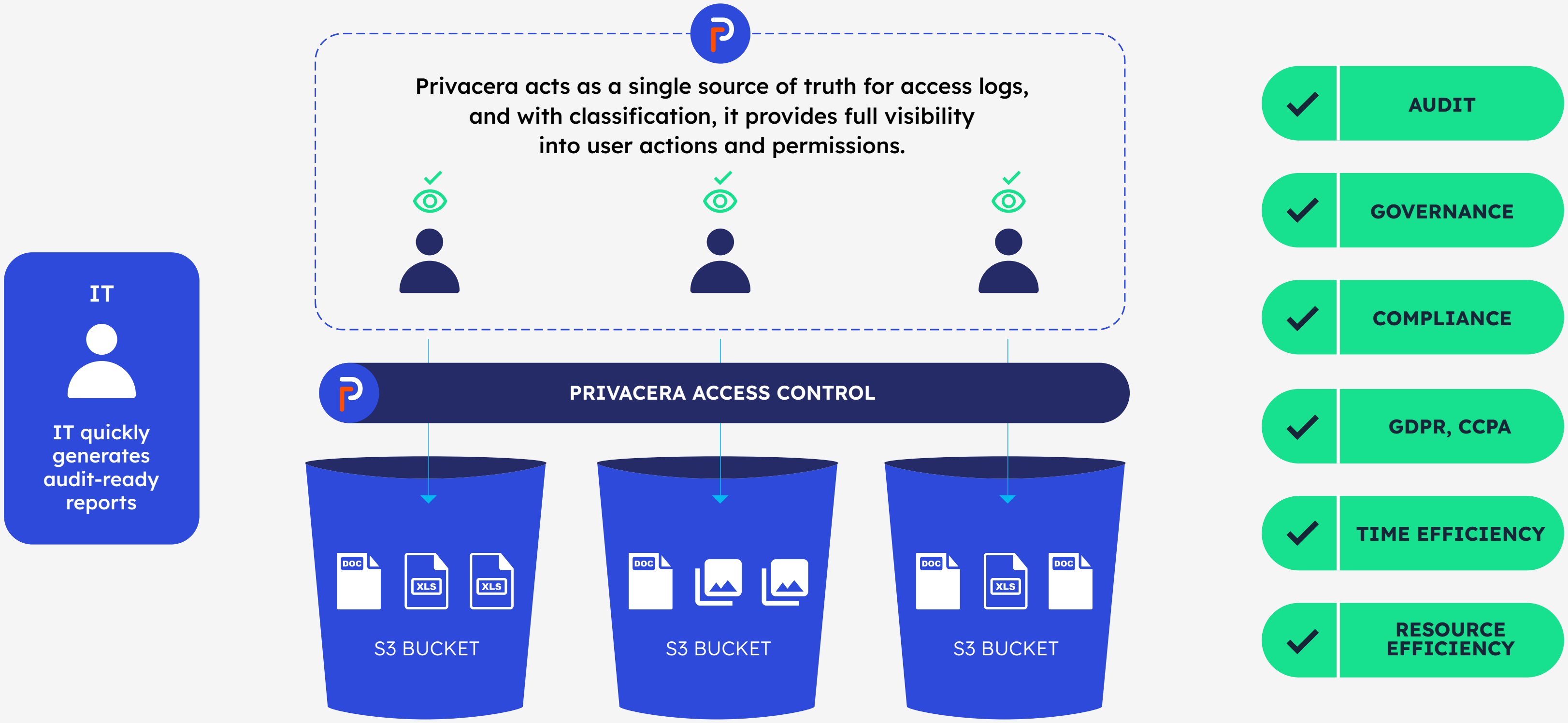
Privacera provides comprehensive visibility of access and governance

With Privacera, you get a complete visibility into access permissions, user actions, and policy enforcement across AWS S3. Privacera enables the real-time enforcement of governance policies, ensuring that users only have access to what they're authorized to see. It does so by acting like a single source of truth for access logs. This enables your organization to quickly generate audit-ready reports to demonstrate compliance. Furthermore, data stewards can classify or tag sensitive data, and policies can be created to provide access based on those classification. This classification helps the compliance managers gain additional visibility into who is accessing sensitive data. As a result, you can gain better oversight and easily meet audits needs.

IDEAL SCENARIO EXAMPLE #5:

Enhancing Retailer Compliance: How Privacera Streamlines Audit and Data Governance

The IT team at Tarmark uses Privacera to manage audit and compliance. Privacera helps them ensure real-time enforcement of access policies across AWS S3, allowing employees to only view authorized data. As Privacera acts as a single source of truth for access logs, it provides full visibility into user actions and permissions. They also classified sensitive data and created polices to provide access based on those classification. This enables the IT team to quickly generate audit-ready reports, ensuring compliance with regulations like GDPR and PCI-DSS. By automating access controls and monitoring, Privacera helps the Tarmark reduce risks, maintain secure customer data, and streamline the audit process.


IT
IT quickly generates audit-ready reports
Privacera acts as a single source of truth for access logs, and with classification, it provides full visibility into user actions and permissions.
PRIVACERA ACCESS CONTROL
DOC
XLS
XLS
S3 BUCKET
DOC
S3 BUCKET
DOC
XLS
DOC
S3 BUCKET
✓ AUDIT
✓ GOVERNANCE
✓ COMPLIANCE
✓ GDPR, CCPA
✓ TIME EFFICIENCY
✓ RESOURCE EFFICIENCY

CHALLENGE #6

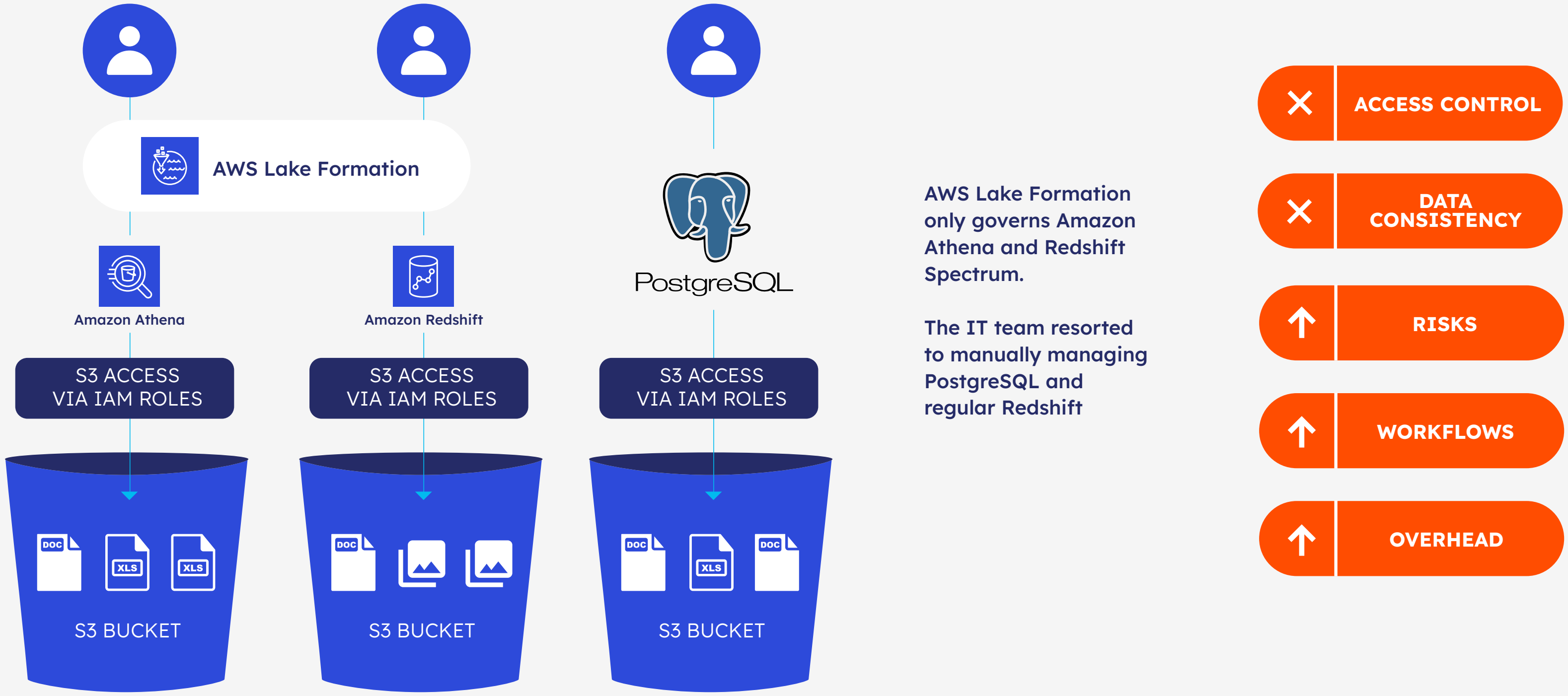
Overcoming Cross-Platform Governance Challenges

To manage cross-platform, many organizations use AWS Lake Formation. However, Lake Formation can only support a limited set of queries, such as those from Redshift Spectrum and AWS Athena. For other data sources, you have to manually manage it. The result of the lack of cross-platform governance increases your effort to manage queries spanning multiple services (Postgres, regular Redshift and other AWS data processing engines). This also introduces the risk of having inconsistencies in access control and policy enforcement across the AWS ecosystem.

UNWANTED SCENARIO EXAMPLE #6:

Integrating Athena, Redshift, and PostgreSQL: Governance Challenges in a Finance Company

In a fintech company, Bank of Narnia, the Data Analytics department uses Athena to query datasets in S3, such as financial history and customer insights. The Transaction Monitoring department relies on PostgreSQL on RDS for real-time financial transaction data. The IT team stores structured data like customer profiles and financial transactions in Redshift, and the Data Integration team uses Redshift Spectrum to query external datasets in S3, including market trends and trading data. They leveraged AWS Lake Formation for governance, but were only able to govern AWS AThena and Redshift Spectrum. They had to manually manage PostgreSQL, and regular Redshift because they are not supported by Lake Formation. The problem was this resulted in inconsistent access control, complex data workflows, and increased overhead. This lack of cross-platform governance raises the risk of unauthorized access.



AWS Lake Formation only governs Amazon Athena and Redshift Spectrum.

The IT team resorted to manually managing PostgreSQL and regular Redshift

SOLUTION #6

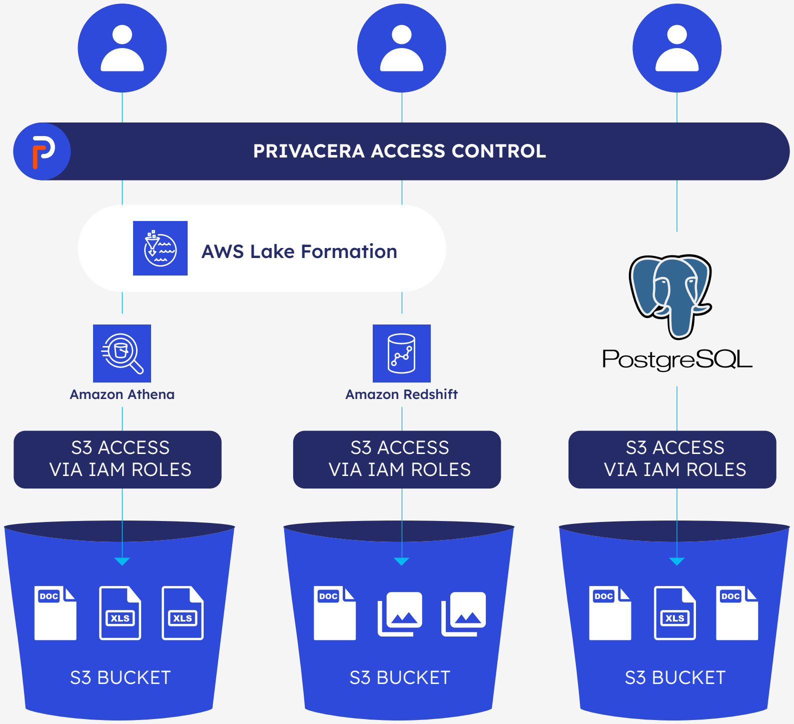
Privacera unifies policy control across diverse services

Unlike Lake Formation, Privacera offers cross-platform governance that supports queries from Redshift, Athena, ProgresSQL, and other AWS services. With Privacera's built-in integrations, access policies are applied consistently across diverse data sources. You have a single, unified control plane for managing AWS resources. Furthermore, Privacera applies consistent policies regardless of the query engine used, ensuring data access control works across multiple AWS analytics services like Redshift, Glue, and Athena.

IDEAL SCENARIO EXAMPLE #6:

Streamlining Data Governance in Finance: How Privacera Solved Cross-Platform Challenges in a Financial Institute

Bank of Narnia implemented Privacera, which helped them gain cross-platform governance by applying consistent access policies across all AWS services, including Athena, Redshift, and PostgreSQL. With Privacera's unified control plane, the IT team could centrally manage access policies, ensuring secure and compliant data access across diverse query engines. This streamlined governance, reduced administrative overhead, and improved data security and compliance across the organization.



With Privacera's unified control plane, the IT team could centrally manage access policies, ensuring secure and compliant data access across diverse query engines

- ✓ GOVERNANCE
- ✓ COMPLIANCE
- ✓ DATA SECURITY
- ✓ ACCESS CONTROL
- ✓ DATA INTEGRITY
- ✓ NO OVERHEAD

IAM	Privacera
Over-permissioning with IAM roles imposes a security risk	**Centralized ABAC with object-level controls provides fine-grained access control**
Creating separate buckets for sensitive data results in IAM role sprawl & complexity (limited to 5000 roles)	**Eliminate role sprawl with object-level control and automated permissions**
Manually granting access control hampers scalability and efficiency (character limit of 2000 to create policies)	**Automated data access, goverance, and compliance increases efficiency** (classification and tagging, and policy)
Manually adding new dataset is time-consuming and error-prone	**Predefined policies instantly gives “birth-right access” to user of the data**
Lack of visibility and auditability of access and permissions (complex, hard to decipher etc)	**Privacera provides comprehensive visibility of access and governance** (centralize platform w/easy to manage policies)
Cross-Platform Governance Challenge	**Privacera unifies policy control across diverse AWS services**

PART 2

Disadvantages of AWS Specific DIY Controls

Organizations often adopt an AWS-specific approach to access and security management, assuming it provides tighter control and cost efficiency. AWS itself promotes its native tools as comprehensive solutions for data governance and security. However, as data volumes grow and extend across various AWS services, this strategy quickly shows its limits. Maintaining custom integrations, managing fragmented access controls, and ensuring consistent compliance across siloed systems becomes increasingly complex and resource-intensive. Without a unified approach, businesses face mounting inefficiencies, security risks, and scalability roadblocks that hinder innovation and strain resources.

DIY Security and Governance Challenges in Hybrid and Multi-Cloud Environments

DIY data governance in hybrid or multi-cloud environments presents two problematic options: managing each data silo separately or stitching together fragmented vendor controls. Both are inefficient and create security gaps. Managing data silos individually leads to inconsistent policies, compliance issues, and higher IT costs. Without standardized frameworks, auditing becomes fragmented, increasing risks and inefficiencies. On the other hand, using cloud vendor solutions may seem cost-effective but results in tech debt and integration challenges, slowing down authorized users and stifling business agility. A unified approach is essential to navigate these complexities and improve operational efficiency.

Operational Impact of Making and Managing DIY on Your Own

Managing cloud complexity on your own is daunting. From classifying metadata to syncing policies and building custom integrations, the cost of a DIY approach is not only in time and resources but in long-term sustainability. Many organizations struggle with the complexity of data integrations and scripts. Some organizations have tried using Google Sheets and custom scripts to manage policies, but they quickly realize that it becomes an unsustainable burden. DIY solutions lack transparency, slow operations, and make compliance harder. Employees face delays in data access and onboarding, and identifying data owners becomes time-consuming, further hindering progress.

Limitation of Integrating DIY with Different Cloud Environments

Cloud vendor solutions may seem cost-effective, but lack scalability, flexibility and integration for hybrid and multi-cloud environments. Each of these vendors operate as a walled garden, forcing enterprises to stitch together controls, creating tech debt that drains resources and hinders agility. This disjointed approach complicates legitimate data access, slowing down authorized users with manual processes, ultimately hindering productivity and operational efficiency. Furthermore, organizations depend on legacy governance methods, which end up becoming roadblocks and delay access to critical data.

Cost in-efficiency with DIY

DIY cloud management comes with tangible costs that many organizations underestimate until they are deep into their cloud journey. Managing the complexity of varied data sources, each with different standards, drives up integration challenges and total cost of ownership.

Cloud is not inherently simple, and without a cohesive strategy, the costs - both financial and operational - quickly escalate. The issues faced by our customers with DIY security, particularly in data management and governance, can be summarized as follows:

Scalability Challenges with DIY

As data volumes grow, manual governance processes become impractical for customers. For example, a financial services institution (FSI) faced a scalability wall with their 31 petabytes of data spread across multiple systems (Redshift, Spark, EMR, Flink, etc.) created complexity and management sprawl. Managing large amounts of data across thousands of tables and datasets proved overwhelming in the end, making it hard to govern, control, and secure access.

Access Control Issues in DIY

Coarse-grained access controls lead to overly permissive and difficult-to-manage access. The inability to efficiently track and manage who had access to what data created significant security risks. The complexity of managing sensitive data in highly regulated industries required dynamic access management solutions to better handle user attributes and group memberships.

Operation Inefficiencies with DIY

For many customers, manually processing access requests (via ticketing systems) and manual onboarding into governance frameworks created operational bottlenecks and error-prone processes. Manual encryption for data masking was inefficient at scale, slowing down data access and adding costs.

Data Visibility and Redundancy in DIY

Some customers lacked central visibility into their data lakes, which hindered the ability to track datasets, distinguish environments, and monitor access. Multiple copies of datasets led to increased storage costs and difficulties in identifying production vs. test datasets.

Redundant and untracked datasets exacerbated costs, while reliance on tribal knowledge for periodic cleanups was unsustainable.

Technology Sprawl and Integration Complexities with DIY

The integration of various technologies like Databricks, EMR, Redshift, and others introduced maintenance and administrative burdens. Ensuring seamless data access across these platforms was difficult. Organizations struggled to manage data governance across multi-cloud environments, further complicating compliance and security efforts.

Governance and Compliance Risks with DIY

Relying on federated permission management (with lower-level controls) without centralized oversight made it hard to ensure proper governance and compliance with regulations.

There were risks of non-compliance due to potential delays in processing data access requests and manual governance approaches. In summary, organizations face difficulties with scalability, access control, operational inefficiencies, data visibility, and governance across complex cloud ecosystems, leading to rising costs, security risks, and challenges with regulatory compliance.

PART 3

How Privacera Can Help

Privacera empowers organizations with comprehensive visibility into all sensitive data across various source systems. With its built-in automation and a robust suite of over 50 connectors developed through extensive engineering efforts, Privacera significantly reduces the need for staff such as data platform administrators and data engineers. Each data platform follows different standards for access policies, creating a heavy burden on IT and security teams to maintain consistent organizational policies.

Privacera automatically detects sensitive data across multiple cloud databases and analytics platforms, allowing rules to be written once and applied across all sources. By leveraging pattern recognition and machine learning, Privacera enhances the discovery and tagging of unprotected data and personally identifiable information (PII). This enables data stewards to make informed decisions regarding what data needs protection and who should have access across various systems.

Moreover, Privacera automates the scanning, identification, and tagging of sensitive data both on-premises and in the cloud, covering the entire data estate. It provides a unified view of access policies, reducing inconsistencies, redundancies, and manual errors associated with policy administration. With fine-grained policies in place, organizations can mitigate the risks of data breaches and leaks, resulting in a 50-75% reduction in the resources required for policy administration through automation.

Greater access to data leads to more informed analytics, reduced time to insights, and faster decision-making, empowering organizations to act swiftly in a data-driven world. By streamlining data governance, Privacera ensures that organizations can navigate the complexities of data security with ease and confidence.

Unified Data Security Platform Advantage

A unified data security platform offers significant advantages akin to the efficiencies gained from standardization in enterprise architecture, as highlighted in “Enterprise Architecture as Strategy.” By adopting a centralized approach to data security, organizations can reduce complexity and minimize the number of platforms they operate, leading to lower costs and IT budgets that are 15% leaner.

This unified governance model delivers immense value by significantly lowering the administrative burden associated with managing data security and access control. Organizations can reduce the number of dedicated resources focused solely on policy administration and data access management, allowing teams to redirect their efforts toward more strategic initiatives. This streamlined approach alleviates the workload on policy administrators while accelerating user onboarding and data access, ensuring that critical business information is readily available.

By consolidating data governance policies under one system, organizations enhance transparency, consistency, and auditability, which improves compliance standards and bolsters security postures. The reduction of manual processes and duplicative controls makes it easier to audit and enforce compliance, enabling employees to gain expedited access to data. This facilitates better decision-making while ensuring regulatory requirements are met.

User/Subject Attributes	Resource/Object Attributes	Environmental Attributes	Action Attributes
✓ Username	✓ Type	✓ Location	✓ Read
✓ Employee ID	✓ Author/Owner	✓ Time Zone	✓ Write
✓ Job Title	✓ Classification	✓ Current Time	✓ View
✓ Department	✓ Date Created	✓ Current Day	✓ Transfer
✓ Clearance	✓ Last Updated	✓ Device	✓ Delete

Moreover, enhanced efficiency in managing data governance not only improves internal processes but also elevates employee satisfaction and customer experiences. With quicker provisioning of data access and reduced labor in managing security, organizations become more agile, better equipped to handle growth and ensure data democratization. This unified approach ensures faster access to data, stronger compliance, and reduced operational complexity, resulting in a more agile and secure environment.

Finally, a unified platform future-proofs your data estate. As new data technologies emerge - often without the input of data teams - the ability to seamlessly integrate new data sources into an existing central platform offers substantial benefits. For instance, one manufacturer/retailer we work with was able to reduce the introduction of new data by up to 95%, demonstrating the transformative power of a centralized data security approach.

Conclusion

In today's complex hybrid and multi-cloud landscapes, CIOs and CISOs must confront the intricate challenges of data governance that traditional enterprise data warehouses (EDWs) no longer address. While cloud vendors like AWS, Databricks, and Snowflake promise solutions, their offerings often fail to integrate effectively across different platforms, leaving organizations struggling with fragmented controls and inefficiencies.

The DIY approach to data governance presents a false sense of control, forcing teams to either manage each data silo independently or piece together disjointed vendor tools, both of which lead to increased operational costs and compliance risks. Without a cohesive strategy and centralized governance framework, organizations face spiraling costs and diminished agility, hampering their ability to respond swiftly to data needs. A unified data security platform emerges as the key to overcoming these challenges, streamlining processes, enhancing compliance, and significantly reducing the administrative burden associated with managing diverse data sources. By leveraging Privacera, organizations can automate governance, gain comprehensive visibility into sensitive data, and ensure that data access is both efficient and secure, ultimately transforming their approach to data security in an increasingly complex environment.

AUTHORS:

Balaji Ganeshan, CEO & Co-Founder, Privacera
Balaji Ganesan is CEO and co-founder of Privacera. Before Privacera, Balaji and Privacera co-founder Don Bosco Durai, also founded XA Secure. XA Secure's was acquired by Hortonworks, who contributed the product to the Apache Software Foundation and rebranded as Apache Ranger.

Don Bosco Durai, CTO & Co-Founder, Privacera
Don Bosco Durai (Bosco) is CTO and co-founder of Privacera, an entrepreneur and a thought leader in enterprise security. He is the co-creator of Apache Ranger, which is the de facto centralized authorization tool for most open source big data tools. Prior to founding Privacera, Bosco was also the co-founder of XA Secure, which redefined access security at scale.

Ibrahim "Ibby" Rahmani
Sr. Director of Marketing & Product Marketing
Ibby Rahmani is the marketing lead at Privacera, with a proven track record of launching successful programs at Alation, Hitachi, HP, and VMware. He specializes in AI, data technologies, and go-to-market strategies, translating complex innovations into clear, impactful messaging. Passionate about driving industry awareness, he creates compelling content that highlights emerging trends and real-world applications.

Jason Payne, Head of Solutions Engineering, Privacera
Jason Payne is the Head of Solutions Engineering and Technical Services at Privacera, specializing in data security and governance. He leads technical strategies for data access and privacy and shares insights through Privacera's technical video series.

Fortune 500 enterprises trust Privacera for their universal data security, access control, and governance. Discover how to streamline data security governance with Privacera.

Take a unified approach to data access, privacy, and security with Privacera.

REQUEST A DEMO →   CONTACT US →

Privacera, based in Fremont, CA, was founded in 2016 by the creators of Apache Ranger™ and Apache Atlas. Delivering trusted and timely access to data consumers, Privacera provides data privacy, security, and governance through its SaaS-based unified data security platform. Privacera's latest innovation, Privacera AI Governance (PAIG), is the industry's first AI data security governance solution. Privacera serves Fortune 500 clients across finance, insurance, life sciences, retail, media, consumer, and government entities. The company achieved AWS Data and Analytics Competency Status, and partners with and supports leading data sources, including AWS, Snowflake, Databricks, Azure and Google. Privacera is recognized as a leader in the 2025 GigaOm Radar for Data Access Governance. Learn more at privacera.com.