IMMUTA

How Immuta Makes Life Easier for Data Architects

Data architects are facing a triad of challenges in the current state of analytics: the shift to cloud data platforms, the need to support varied workloads across BI and data science, and the increase in sensitive data collection from digital channels for predictive modeling and prescriptive analysis. In today's analytics environment, traditional approaches to access control no longer scale in the cloud.

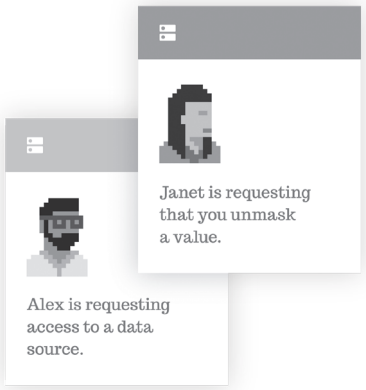
This brief explores the top challenges raised by data architects who have selected Immuta to scale secure access control for their cloud data platforms, such as Amazon Redshift, Azure Synapse, Databricks, Google BigQuery, Snowflake, Starburst, and others. Let's take a closer look at how Immuta can help make your life easier.

Data Architects' Top Cloud Challenges

In the cloud, the concept of "separating storage from compute" means that storage configurations are decoupled from computation resources. This allows applications to scale computing and storage needs independently, but also means that traditional approaches to access control may no longer work as expected. The following data access challenges are compounded when adopting cloud data architectures.

01 Manual steps to provision data for users

The process to fulfill a security exemption for analytics is often manual and can take anywhere from weeks to months. Organizations face high opportunity costs when data cannot be accessed by an increasing number of data consumers (business analysts, technical analysts, and data scientists) whose usage is growing and expected to scale further with cloud data platform adoption. Increasing compute layers on cloud-based object stores introduces disparate approaches to data security and privacy, since approaches to data access policy implementation do not scale and vary greatly by compute technology. This includes creating role-specific views, writing scripts of Data Definition Language (DDL) operations to implement native controls where available, writing compute-specific functions or custom plugins, or replicating data using complex Extract, Load, Transform (ELT) processes.



An example scenario is a Fortune 500 organization's Human Resources (HR) team that used Databricks on Azure to run models to predict resourcing for their remote workforce. HR's existing approach to leveraging sensitive data in the cloud was disrupted when it was blocked by the compliance team. In order to provision access to this data, the data platform team was required to anonymize the data and provide a full auditing trail for every use case deployed. Manually handling these requests was nearly impossible to scale.

Impact of manual provisioning steps to the data platform team:

- Data engineering resources get tied up in manual administrative tasks
- ROI on the data platform product decreases when data consumers are unable to access data in a timely fashion
- Manually provisioning data access increases risk

02 Inability to get necessary auditing detail for compliant data use

Security and regulatory compliance teams ask seemingly simple questions such as: "Who accessed this data over the last 8 weeks?" But answering these questions is very complex in the context of modern cloud platforms.

Working with disparate compute layers, in which security and privacy controls may be enforced in the compute layer or passed through to storage (which is often an object store), makes it difficult to get consistent data auditing and efficient reporting. In other cases, access control rules may be implemented in analytical applications such as Tableau or Power BI. Further, data architects support many data consumers, including data scientists, business analysts, data stewards, etc., who often reside in different business units. All of this can lead to inconsistent policy enforcement that makes monitoring data use complicated and unreliable.

As in the HR example, many data platform teams are required to de-identify, or anonymize, personally identifiable information (PII), which includes a requirement to provide a full auditing or reporting trail for every data interaction. This is often mandated by internal rules or the ever increasing data protection laws such as the CCPA and GDPR. Failing to meet these standards could result in substantial monetary and reputational damages.

Impact to the data platform team from gaps in auditing data use:

- Lack of consistent auditing limits acceptable use cases for cloud analytics
- Increased surface area of risk for improper data use that result in data leaks
- Responses to security or regulatory compliance requests that are manual or not even possible in the current architecture

IMMUTA

How Immuta Makes Life Easier for Data Architects

Data architects are facing a triad of challenges in the current state of analytics: the shift to cloud data platforms, the need to support varied workloads across BI and data science, and the increase in sensitive data collection from digital channels for predictive modeling and prescriptive analysis. In today's analytics environment, traditional approaches to access control no longer scale in the cloud.

This brief explores the top challenges raised by data architects who have selected Immuta to scale secure access control for their cloud data platforms, such as Amazon Redshift, Azure Synapse, Databricks, Google BigQuery, Snowflake, Starburst, and others. Let's take a closer look at how Immuta can help make your life easier.

Data Architects' Top Cloud Challenges

In the cloud, the concept of "separating storage from compute" means that storage configurations are decoupled from computation resources. This allows applications to scale computing and storage needs independently, but also means that traditional approaches to access control may no longer work as expected. The following data access challenges are compounded when adopting cloud data architectures.

01 Manual steps to provision data for users

The process to fulfill a security exemption for analytics is often manual and can take anywhere from weeks to months. Organizations face high opportunity costs when data cannot be accessed by an increasing number of data consumers (business analysts, technical analysts, and data scientists) whose usage is growing and expected to scale further with cloud data platform adoption. Increasing compute layers on cloud-based object stores introduces disparate approaches to data security and privacy, since approaches to data access policy implementation do not scale and vary greatly by compute technology. This includes creating role-specific views, writing scripts of Data Definition Language (DDL) operations to implement native controls where available, writing compute-specific functions or custom plugins, or replicating data using complex Extract, Load, Transform (ELT) processes.

An example scenario is a Fortune 500 organization's Human Resources (HR) team that used Databricks on Azure to run models to predict resourcing for their remote workforce. HR's existing approach to leveraging sensitive data in the cloud was disrupted when it was blocked by the compliance team. In order to provision access to this data, the data platform team was required to anonymize the data and provide a full auditing trail for every use case deployed. Manually handling these requests was nearly impossible to scale.

Impact of manual provisioning steps to the data platform team:

- Data engineering resources get tied up in manual administrative tasks
- ROI on the data platform product decreases when data consumers are unable to access data in a timely fashion
- Manually provisioning data access increases risk

02 Inability to get necessary auditing detail for compliant data use

Security and regulatory compliance teams ask seemingly simple questions such as: "Who accessed this data over the last 8 weeks?" But answering these questions is very complex in the context of modern cloud platforms.

Working with disparate compute layers, in which security and privacy controls may be enforced in the compute layer or passed through to storage (which is often an object store), makes it difficult to get consistent data auditing and efficient reporting. In other cases, access control rules may be implemented in analytical applications such as Tableau or Power BI. Further, data architects support many data consumers, including data scientists, business analysts, data stewards, etc., who often reside in different business units. All of this can lead to inconsistent policy enforcement that makes monitoring data use complicated and unreliable.

As in the HR example, many data platform teams are required to de-identify, or anonymize, personally identifiable information (PII), which includes a requirement to provide a full auditing or reporting trail for every data interaction. This is often mandated by internal rules or the ever increasing data protection laws such as the CCPA and GDPR. Failing to meet these standards could result in substantial monetary and reputational damages.

Impact to the data platform team from gaps in auditing data use:

- Lack of consistent auditing limits acceptable use cases for cloud analytics
- Increased surface area of risk for improper data use that result in data leaks
- Responses to security or regulatory compliance requests that are manual or not even possible in the current architecture

03 Lack of data security to migrate or consolidate cloud analytics

Cloud data platforms are still in nascent stages of maturity when it comes to data security and privacy. As cloud migration accelerates, data architects face a proliferation of rules that need to be implemented to process sensitive data. However, traditional access policies were often built using tools for static data masking or tokenization that were engineered for on-premises data sets with limited user roles.

These approaches no longer scale in the cloud and are driven by security or regulatory compliance teams in response to ever-changing rules around data use from data protection laws, data sharing agreements, employment laws, intellectual property controls, SOC-2 compliance effort, etc. The data team, in turn, receives a list of requirements (see figure below for example of interpretation) that they must enforce accordingly. These rules are bound to disrupt platform operations given the current state of data security and governance in cloud analytics.

For example, in the CCPA example below, the rules are given to the data team as requirements to implement, and include determining and proving that PII data is properly de-identified for use, imposing purpose-based restrictions for data use (leading to unscalable manual processes), or automating data retention policies for different classes of users.

Impact to the data platform team from nascent cloud data security:

- Lack of consistent auditing limits acceptable use cases for cloud analytics
- Increased surface area of risk for improper data use that result in data leaks
- Responses to security or regulatory compliance requests that are manual or not even possible in the current architecture

Sample Language from CCPA Exclusions, Opt Out, and Purpose

LEGAL/COMPLIANCE INTERPRETS	DATA TEAM IMPLEMENTS
1798.145(a)(5) – Exclusions for de-identified & aggregated data	Build data policies to obtain these exclusions that properly de-identify the data with mathematical guarantees that prevent re-identification
1798.100 & 1798.140(w) – Purpose Restrictions Internally and for Vendors	Build and manage secure data "zones" by purpose Develop consent workflow to acknowledge intended use
1798.120(a) – The Right to Opt-Out	Purge records or create row-level policies to hide data based on identifier(s) to comply with opt-out requests

HOW CAN IMMUTA MAKE YOUR LIFE EASIER?

Automate data access to provision data to users in minutes, rather than months

Traditional data access control strategies require manual request workflows combined with role-based access control (RBAC), which were formalized 30 years ago for relational databases.

Given the significant changes to architecture and data protection requirements over that period of time, data engineering and administration teams suffer from "role explosion" as a result of mapping hundreds or thousands of policies to user roles that control fine-grained access to data.

To address this challenge, Immuta leverages attribute-based access control (ABAC), which uses dynamic attributes to enforce data protection at query time. Because Immuta's architecture decouples users and authorization, it can enforce policies based on metadata and user attributes such as geography, time and date, clearance level, and purpose. A single Immuta ABAC policy enables self-service and authorized access, and can replace over 100 roles, saving time and reducing security risks.

GOAL	RBAC	ABAC
Flexibility	Small-sized orgs	✓ Yes
Scalability	✗ No	✓ Yes
Simplicity	Easy start, hard to maintain	Work up front, easy to maintain
Simple rules	✓ Yes	✓ Yes
Complex rules	✗ No	✓ Yes
Dynamic rules	✗ No	✓ Yes
Customizing permissions	Requires new roles	✓ Yes
High granularity	✗ No	✓ Yes
Policy comprehension	✗ No	✓ Yes

HOW CAN IMMUTA MAKE YOUR LIFE EASIER?

Implement a centralized audit plane

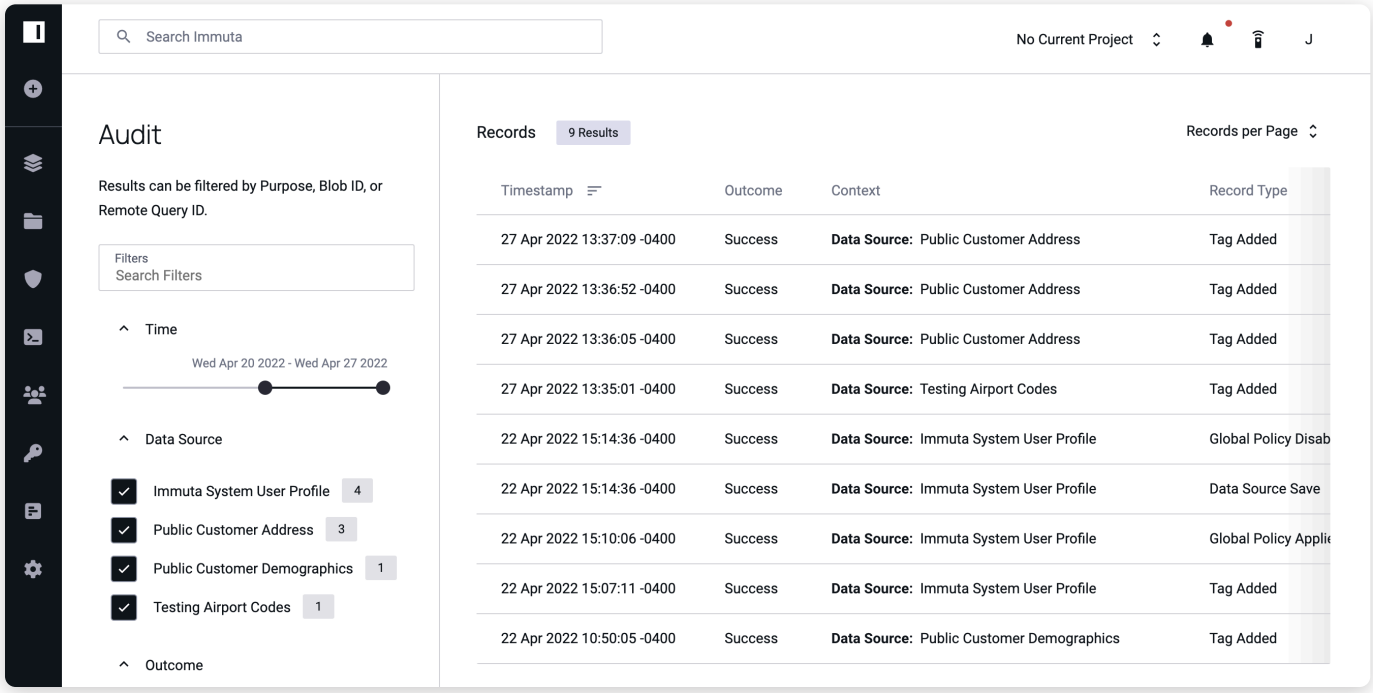
With Immuta, you can monitor and log all actions in the Immuta control plane directly from your cloud data platform to prove compliant data use. This enables you to track requests and access to data, policy changes, data usage reports, precise queries executed by users, and more, all from a centralized policy dashboard.

Data platform teams use Immuta to automate fulfillment of compliance teams' requests to understand which users have accessed a specific data source, or the purposes for which a data set has been used. Immuta reports are built to quickly answer these questions – and virtually any other query – from your compliance team. This empowers more team members to understand and manage data policies, thereby enabling distributed stewardship.

In addition to auditing data use across disparate layers in the modern data stack, data platform teams must make data use simple and transparent for a growing number of stakeholders.

User-friendly UIs are becoming a requirement, particularly for non-technical stakeholders who need to understand how data is being provisioned and used, but who may not be well-versed in native controls or basic metrics collected by cloud providers.

Using Immuta's explainable policy builder, data teams are able to create policies in plain language so that all legal, regulatory compliance, and governance stakeholders can understand how analytics access control is managed and protected.



Learn more about Immuta's audit logs and reports.

HOW CAN IMMUTA MAKE YOUR LIFE EASIER?

Manage Security and Privacy Controls without Data Replication

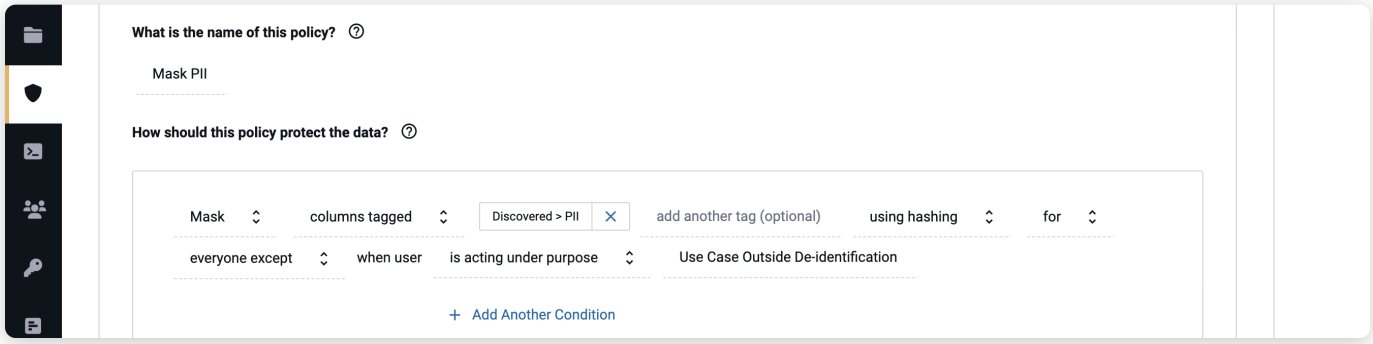
The modern approach to scaling security and privacy controls in the cloud is to apply policies at query time through dynamic data masking, without having to replicate underlying data, as would be the case with static data masking.

Immuta provides an explainable policy builder to author dynamic data masking and protection policies, empowering data teams to deliver consistent data security and privacy controls across data platforms. Policies can be enforced locally on a single table, or globally across the data platform. This removes the guesswork from implementing ever-changing data protection requirements driven upstream by legal, compliance, or governance teams. Using the dynamic data masking approach, authored policies are applied at query time and do not require creating views or ETL processes to transform and replicate data, which may diminish its utility.

Immuta's masking policies support the latest privacy enhancing technologies (PETs) to securely satisfy novel requests to protect data while preserving utility. These technologies include hashing, regular expression, rounding, conditional masking, k-anonymization, and replacing with null or constant, with reversibility, or with format preserving masking, as well as more advanced techniques for randomization, such as differential privacy or randomized response.

These allow policies to handle scenarios across cloud compute layers in a consistent way, such as:

- Restricting access using time-based policies
- Restricting access by geography
- Masking data with the format preserved for non-production use
- Handling unmasking requests when access to raw data is needed, such as for a help desk
- Limiting data use to specific purposes through purpose-based access control (PBAC)
- Protecting against data leaks when users with heightened access create derivative data sets in a workspace
- Preventing linkage attacks using dynamic k-anonymization on individual data
- Enforcing rules for use in data sharing agreements
- Data minimization policies restricted to a percentage of available data
- Manage granular access control to intellectual property
- And hundreds more ...



Learn more about subscription policies and data policies in Immuta

How Immuta Helps Data Architecture Teams

"We needed to expedite our data processing, while also finding a way to dynamically anonymize sensitive information for reporting. We therefore required a solution that could help us enforce data access roles, permissions and policies beyond the standard resource- or table-based control levels."

— **Halim Abbas**, Chief Ai Officer, Cognoa

"Databricks gives us scale and speed, Immuta gives us trust and privacy, Databricks and Immuta together are a good chunk of what we offer to our research team to work with."

— **Slava Frid**, CTO, WorldQuant Predictive

"By incorporating Immuta's automated governance and privacy capabilities, we have enhanced our overall strength and security of the platform. With its Series C funding, we look forward to continued innovation from Immuta to help Aon with our data and analytics initiatives."

— **Steve Petrevski**, SVP and General Manager, Data & Analytics Services at Aon

Why Immuta?

The chart below summarizes the top challenges raised by data architects and how Immuta addresses them, enabling secure and governed access control to effectively scale cloud data analytics.

Traditional Approach to Cloud Data Access & Security	Modern Approach to Cloud Data Access & Security
Request access to sensitive data, which requires manual approvals that take months.	Enable instant, self-service access to sensitive data with rules that are dynamically enforced at query time. Scale across hundreds of users using attribute-based access controls (ABAC) without making any changes to underlying policies.
Build an audit system and UI on top of your data platform based on evolving requirements from compliance teams.	Provide real-time audit logs and reports that demonstrate compliant data use and policies in plain language, without building and maintaining a DIY audit system.
Write code to build and maintain data pipelines that implement security and privacy controls for each user.	Scale your data pipelines by applying dynamic and flexible data policy enforcement at query time, without replicating data or writing code.

While this brief is focused on giving data architects a clear picture of Immuta and its benefits, Immuta is engineered for data platform teams looking to centralize data access platform security and governance. Our data access platform allows data architecture and engineering teams to safely analyze sensitive data in the cloud, and our unique focus on collaboration between data and legal or compliance teams enables organizations to derive greater value from their cloud data platforms.

The Immuta Data Access Platform delivers data access and security at scale for global financial services, healthcare, pharmaceutical and life sciences companies, the public sector, and leading manufacturing, technology, and entertainment brands. Immuta discovers, secures, and monitors an organization's data to ensure that users have access to the right data at the right time – as long as they have the rights.

Ready to get started? Request a demo today.

IMMUTA

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