IMMUTA + databricks

Automate Security and Privacy Controls for Data Science and BI

Immuta for Databricks governs production analytics for the organizations that compete with data yet must preserve privacy – including global banks, healthcare companies, intelligence agencies,technology companies, and consumer brands.

Together with Databricks, Immuta helps data teams automate security and governance when sensitive data is used for BI and data science. The solution simplifies data preparation time, eliminates manual masking and anonymization, and mitigates security risks while delivering subscription-level, self-service access to analysts and data scientists.

What are the Key Challenges?

Row- and Column-Level Security

Managing row- and column- level security controls at scale results in proliferation of views, data copies, and ETL jobs for each user access policy.

Audit Data Access for Compliance

Providing full transparency with compliance and legal teams requires detailed auditing to accurately answer requests for: what data was accessed, by whom, when, and for what purpose.

Catalog for Data Science and BI

Presenting a unified, self-service data catalog to users is challenging when different tools are used for Data Science and BI.

What are Common Use Cases?

Modernize Access for Hadoop Migration

Automate security and privacy controls when migrating from Hadoop, without the limitations of role-based access control and such tools as Apache Ranger.

Expand Use Cases to BI

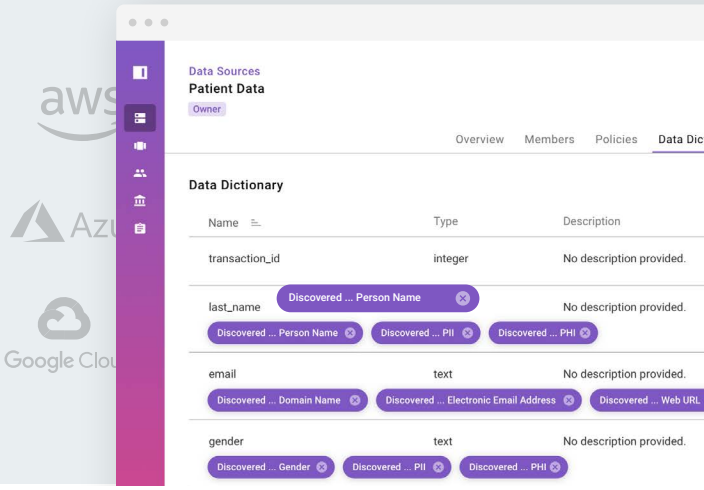
Automate security and privacy controls when migrating from Hadoop, without the limitations of role-based access control and such tools as Apache Ranger.

Access Controls for Data-as-a-Service

Enforce access and privacy controls for each customer of your data product by dynamically enforcing rules natively in Spark without replicating data.

Safely Unlock Sensitive Data with Immuta for Databricks

For teams using Databricks to unify data security and privacy controls for data science and BI, Immuta delivers automated security and privacy controls that enable the safe analysis of sensitive data at scale. Immuta's unique focus on collaboration between data and compliance teams enables organizations to do more with Databricks.



Data Security

Manage data security in sensitive environments with modern, fine-grained, attribute-based access controls (ABAC) dynamically enforced on Spark jobs in Databricks.

Data Governance

Easily prove compliant data access with detailed audit logs and reports that show users' data access levels, intended purposes, and query history – all in plain English.

Unified Data Catalog

Use a self-service data catalog with automated local and global policies applied to Spark workloads in Databricks across Data Science and BI.

RESULTS FOR DATA TEAMS

40% Increase in data engineering productivity when managing sensitive data.

Increase permitted use cases for cloud analytics from

25% - 90%

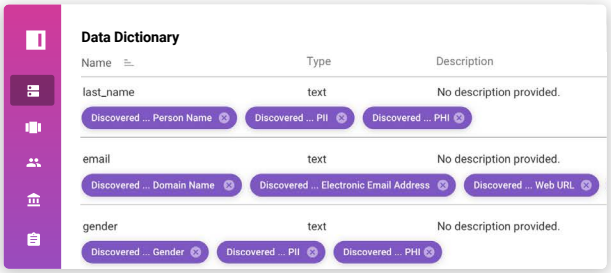
by safely unlocking sensitive data.

Reduce to seconds

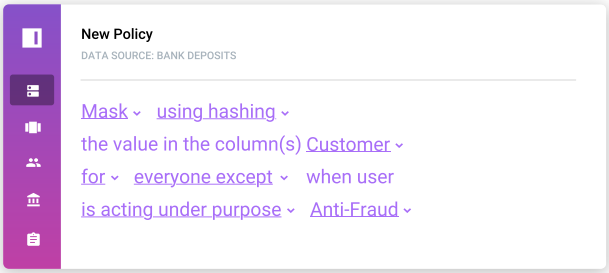
what can be a months-long process to provide self-service data access.

How Immuta for Databricks Works?

01 Catalog, tag and understand your data

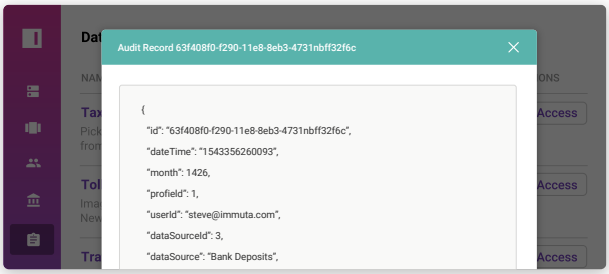


02 Author access control policies using natural language

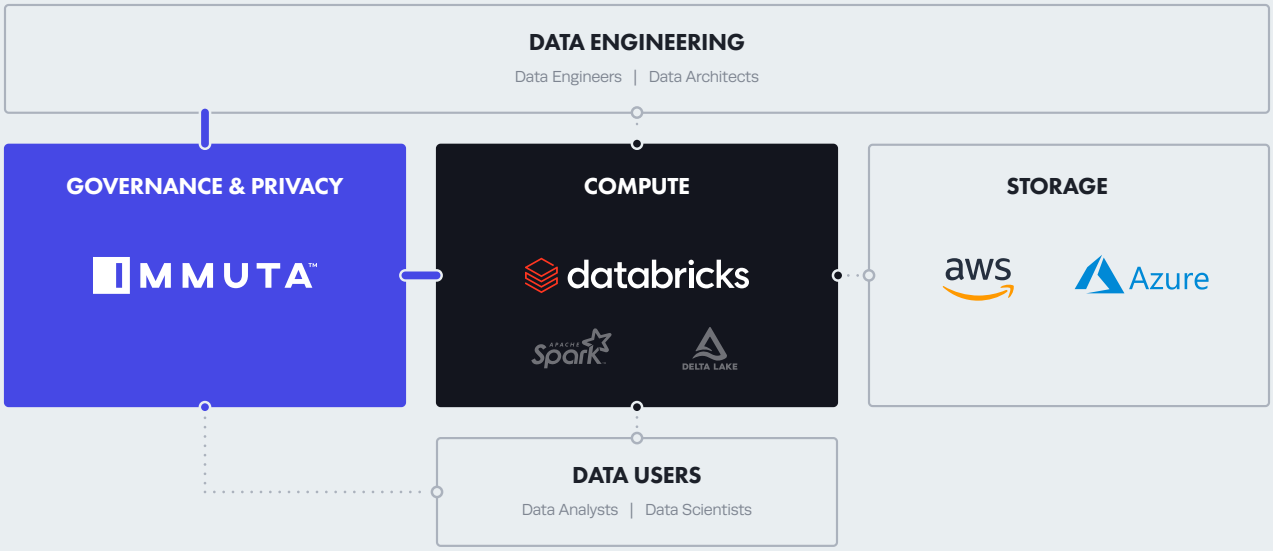


03 Rules are natively enforced in Databricks and transparent to notebooks or BI tools

04 Prove compliance in plain english using detailed audit logs at the data-level



Architecture



Immuta for Databricks Capabilities

	CAPABILITY	DESCRIPTION	BENEFITS
SECURITY	Explainable Policy Builder	Author data policies in plain english that are easy to understand for compliance and legal stakeholders.	
	Row-Level Security	Use Explainable Policy Builder to create dynamic rules to restrict data access on a row-by-row basis to govern what a given user is authorized to see.	
	Column-Level Security	Use Explainable Policy Builder to create dynamic rules to mask data in sensitive columns to govern what a given user is authorized to see.	
	Attribute-Based Access Controls (ABAC)	Use data attributes to write and scale ABAC policies across hundreds of roles, without the limitations of RBAC and such tools as Apache Ranger.	
	Purpose-Based Access Controls (PBAC)	Limit data use to specific purposes with PBAC policies, ensuring that all data use is compliant with data protection laws.	
	R and Scala support for Protected Tables	Enable data scientists to use R, Scala, Python and SQL for both coarse (table-level) and fine (row-, column-, cell-level) grained access controls. Without Immuta, access is limited to coarse grain access with Python and SQL.	
DATA GOVERNANCE	Data Catalog	Provision authorized, self-service user access to sensitive data without manual requests or approvals.	
	Data Collaboration	Create Immuta projects with appropriate access to read and write derivative data sets for safe collaboration for data teams on a Databricks cluster.	
	External Data Sharing	Build data-as-a-service products faster by automating data pipelines using Immuta-powered access controls, enforced natively in Spark using customer attributes.	
	Sensitive Data Discovery	Automatically identify and classify sensitive attributes within your data sets to enforce policies for internal or regulatory compliance at scale.	
	Audit & Compliance Reports	Provide full transparency for legal/compliance teams with detailed logs and reports, at the data level, that show what data was accessed, by whom, when, and for what purpose.	
	Regulatory Starter Policies	Reduce risk of non-compliance with consumer data privacy laws, such as CCPA, HIPAA and others, using global Starter Policies that automate the manual steps of sensitive data discovery, data de-identification and tracking purpose and consent for use.	
PRIVACY	Dynamic Data Masking	Apply data policies to mask, anonymize or randomize data across hundreds of tables – dynamically enforced in Spark jobs, without copying any data.	
	Dynamic Anonymization	Apply advanced anonymization techniques (e.g. k-anonymization) to protect direct and indirect identifiers, as well as sensitive information- dynamically enforced in Spark jobs, without copying any data.	
	Dynamic Randomization	Apply advanced randomization techniques (e.g. differential privacy) to protect direct and indirect identifiers, as well as sensitive information- dynamically enforced in Spark jobs, without copying any data.	

Data Engineer Compliance/Legal Data Consumer (Scientist/Analyst)

14 DAY FREE TRIAL

We invite you to spend 14 days exploring a fully-functional instance of Immuta, for free.

www.immuta.com/try

IMMUTA™ immuta.com (800) 655-0982