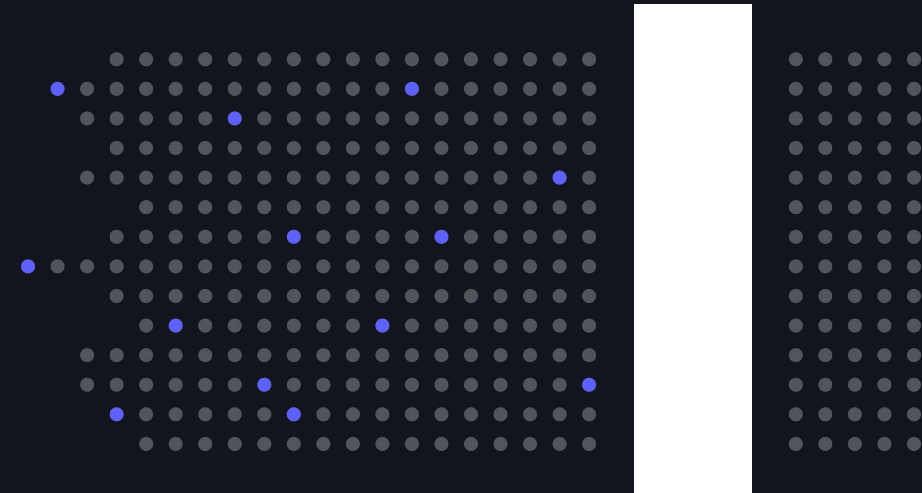




# Dataversity Data Governance Demo Day

**Matt Vogt**

Director, Solution Architecture  
Immuta

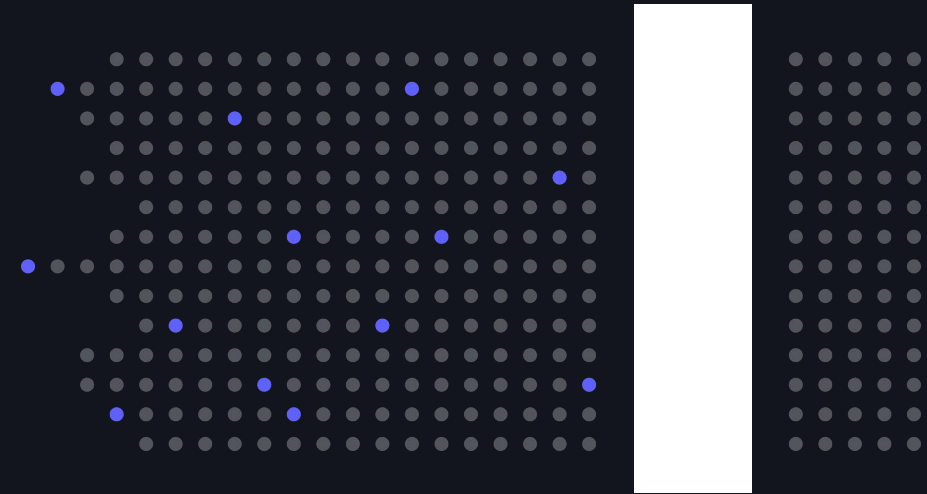




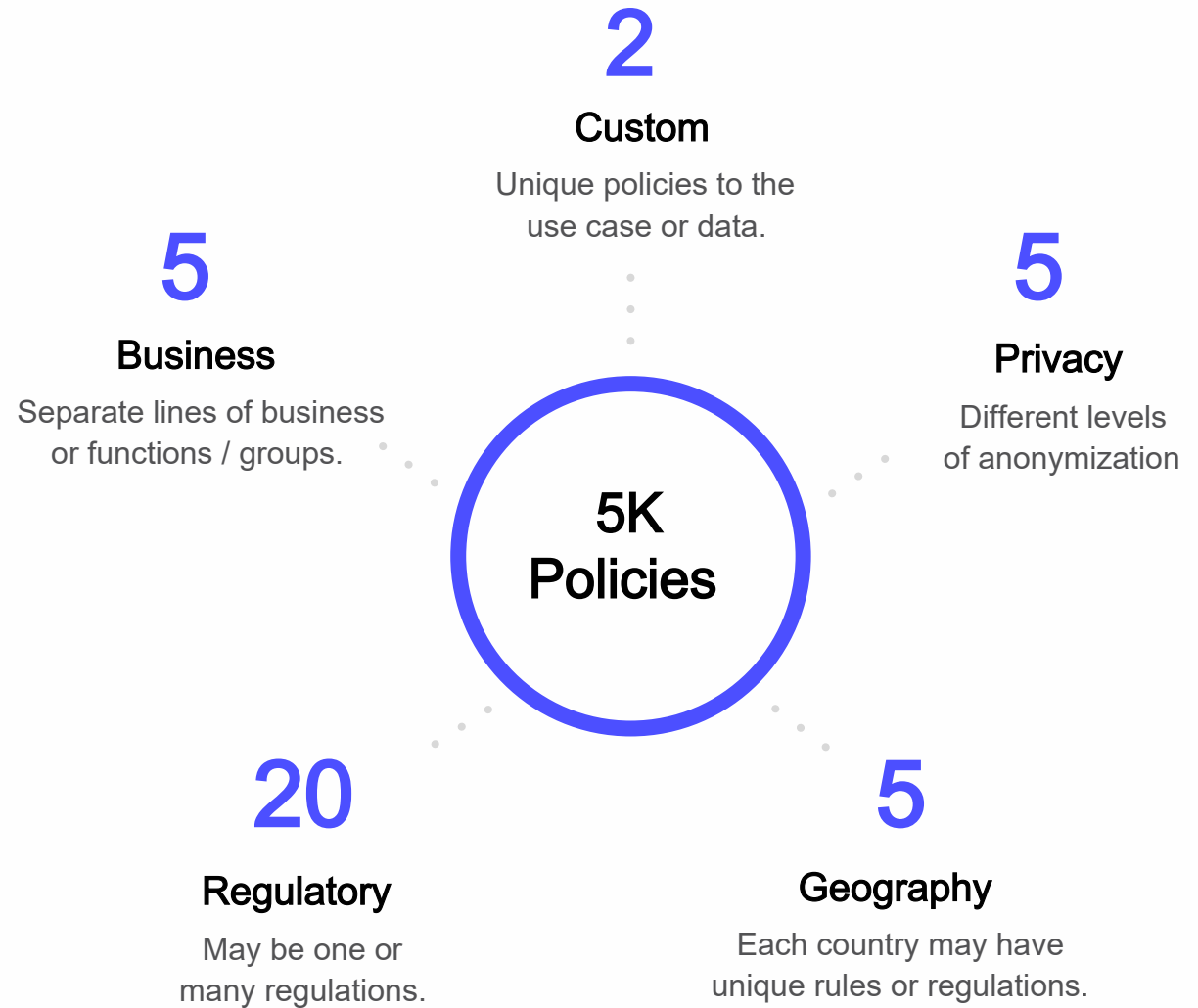
# Introduction to Automated Data Governance

**Matt Vogt**

Director, Solution Architecture  
Immuta

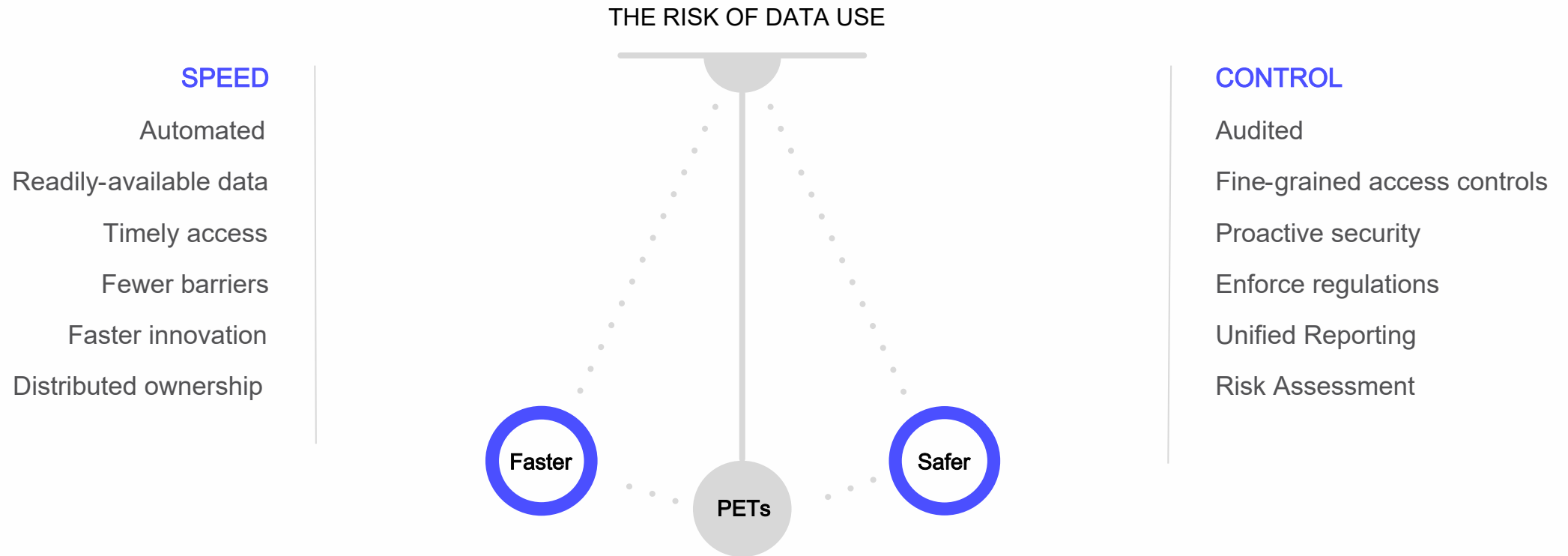


# Trend 1: Complex Privacy Rules are Invalidating the Role Based Access Control (RBAC) model

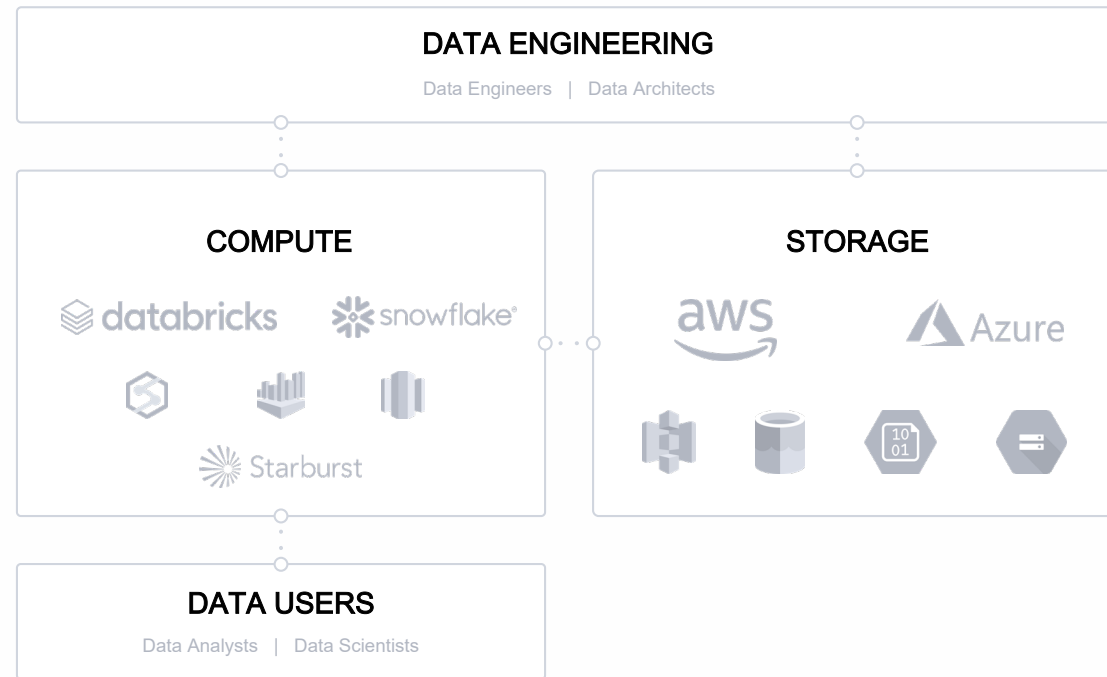


# Trend 2: The Need For Privacy Enhancing Technology (PET)

How to balance the speed of the business with secure access to sensitive data?



# Trend 3: The Separation of Compute & Storage



ABOUT IMMUTA

# Enabling the legal and ethical use of data

We give corporations the ability to take control of their data and enforce natural laws of data – allowing them to do the right thing when processing sensitive data.

- Born in U.S. Intelligence
- 10,000s of users
- Exabyte scale
- Protecting the most sensitive data in the world



**2.5x**  
YoY GROWTH

**3 Offices**  
DC, BOSTON,  
COLUMBUS

TRUSTED BY LEADING  
CLOUD PARTNERS

**2015**  
FOUNDED

RECOGNIZED FOR  
INNOVATION

USED BY COMPANIES  
THAT COMPETE WITH DATA

# Key Features

IMMUTA CAPABILITIES



## SECURITY (ABAC)

---

Row-Level  
Column-Level Security  
Ease-of-Administration



## PRIVACY (PETs)

---

Data Masking  
K-Anonymization  
Differential Privacy



## GOVERNANCE

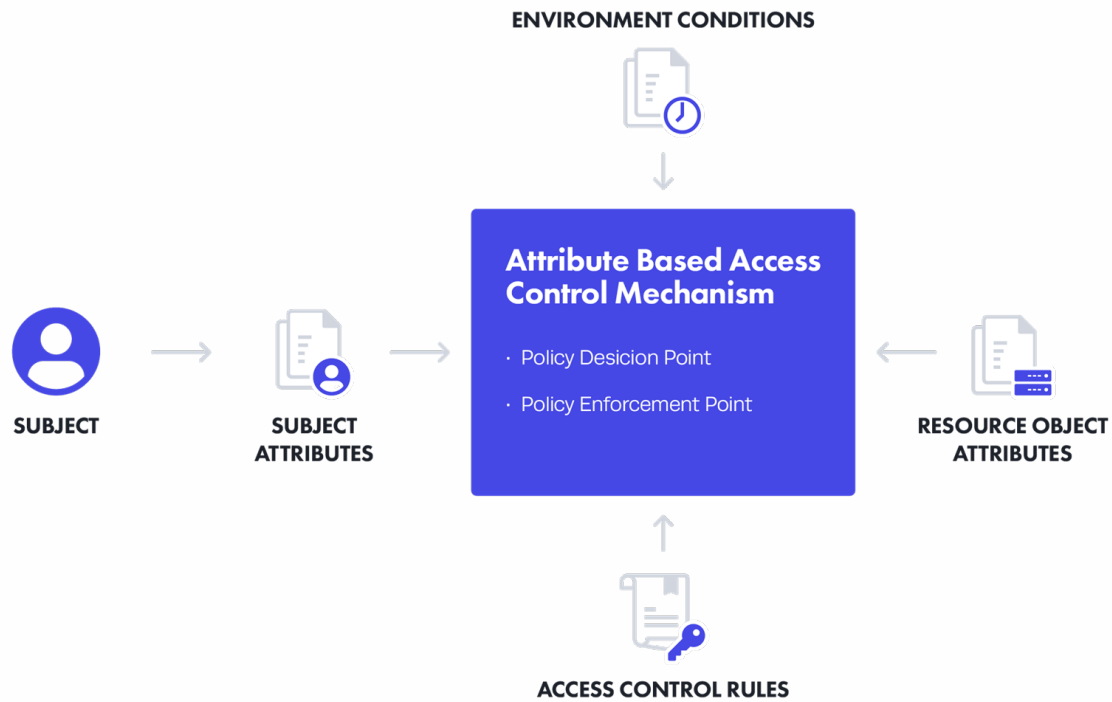
---

Catalog  
Prioritize  
Secure





# Security (ABAC)



Global Subscription Policy Builder

What's the name of this policy?  
[HIPAA Compliance](#)

How should this policy grant access?  
[Click here to select a subscription level to begin building your policy](#)

Where should this policy be applied?  
[On data sources](#) [select condition](#)

- All Circumstances
- tagged
- with columns tagged
- with column names spelled like in server
- created between







# Security (ABAC)

## Granular Control

- Flexible Access Controls
- Attribute-based (ABAC)
- Purpose-based (PBAC)

## Dynamic Policy Enforcement

- Use metadata to drive policies
- Avoid “Role Bloat”
- Distributed administration

## Plain English Rules

- Collaborate IT, LOB, Legal/Compliance on single platform
- Meets any legal framework

Global Subscription Policy Builder

What's the name of this policy?  
[HIPAA Compliance](#)

How should this policy grant access?  
[Click here to select a subscription level to begin building your policy](#)

Where should this policy be applied?  
[On data sources](#) [select condition](#)

- All Circumstances
- tagged
- with columns tagged
- with column names spelled like in server
- created between





# Privacy (PETs)

## Mask

- Author-Once, Protect Everywhere
- Centralize Common Rules
- Advanced Techniques
  - Preserve Format
  - Join on Masked Data

## Anonymize

- k-anonymization
- Rounding
- Differential Privacy

## Audit

- Privacy-by-Design
- Central Auditing

Immuta Demo Project (Python)

Demo ClusterHC

```
Cmd 1
1 %sql
2 SELECT * from immuta.credit_
```

(1) Spark Jobs

id	credit_card_number
1	5602257638494770
2	3052611929219900
3	3539021508894060
4	3529549059441510
5	5602247251428780
6	6371568572188060
7	3540110992660330
8	3537425665907410

Command took 1.72 seconds -- by a user

Immuta Demo Project (Python)

Demo ClusterHC

```
Cmd 1
1 %sql
2 SELECT * from immuta.credit_
```

(1) Spark Jobs

id	credit_card_number
1	*****4770
2	*****9900
3	*****4060
4	*****1510
5	*****8780
6	*****8060
7	*****0330
8	*****7410

Command took 1.72 seconds -- by a user

WITHOUT IMMUTA

WITH IMMUTA

(Dynamic Masking, without copying any data sources)





# Data Governance and Risk

## Catalog

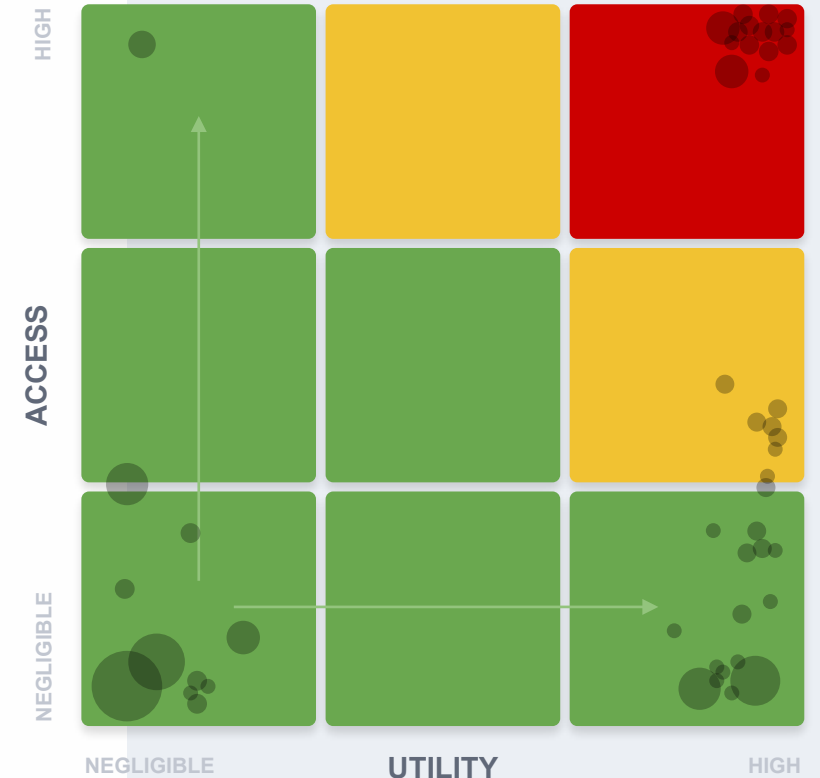
- Inventory data assets and usage
- Analyze Access & Purpose
- Rules Visible & Understandable for all roles

## Prioritize

- Prioritize data sets from under-utilized to over-exposed
- Customize your risk Tolerance
- Retire unused data

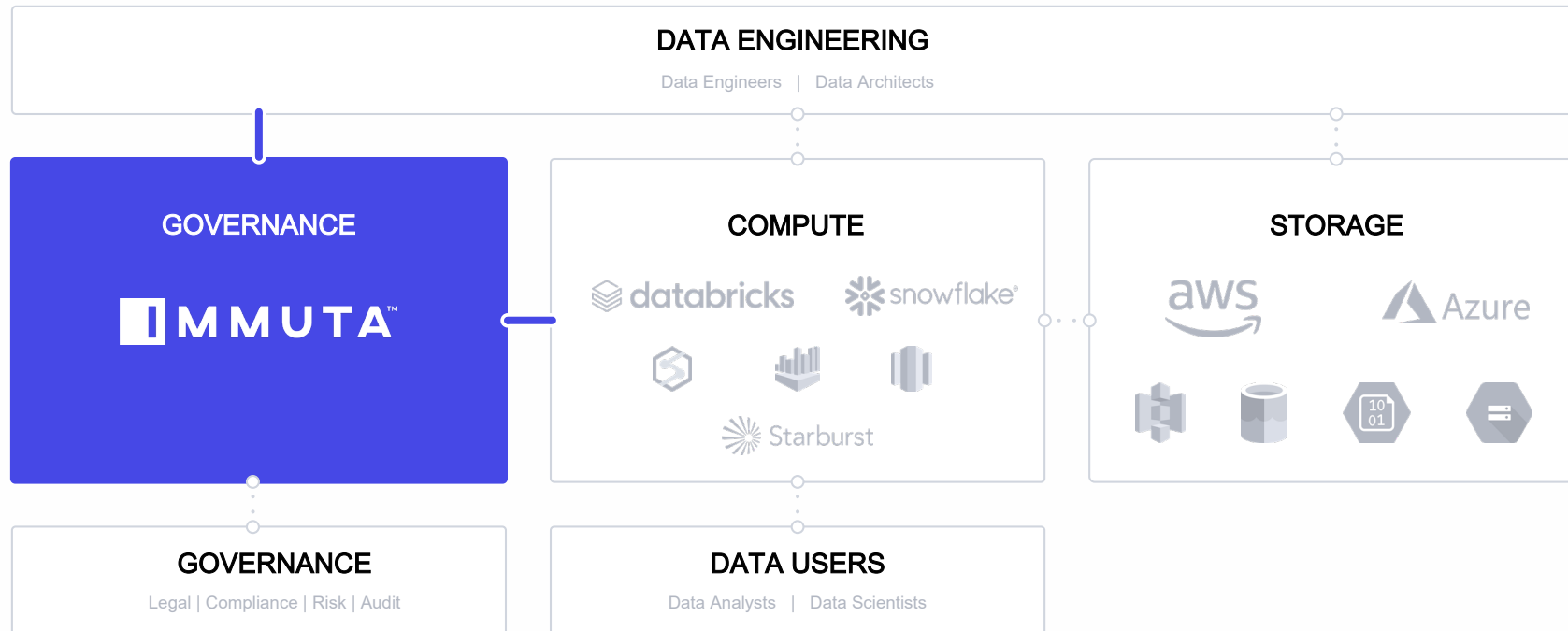
## Secure

- Machine-Learning Policy Creation
- Automated Reporting
- Prove who used what data, when, and why



# Separation of Policy From Platform

The **big data era** required **separation of compute/storage to scale** and now the **personal data era** requires **separation of platform/policy to scale**



# Fortune 500 Transportation Services Company

## BACKGROUND

One of the largest transportation, delivery, and logistics companies in North America.

## TECHNOLOGY & PARTNERS

- Immuta
- Databricks
- Artis Consulting (SI)

## Results:

“With Databricks we were able to accomplish 25% of our total use cases. With Databricks + Immuta, that coverage went up to 90%”

–Artis Consulting

## CASE STUDY

# HR Analytics

## Business Challenge

- **Core Requirement:** The company’s HR team needed to run models to predict resourcing for their fleets. Their compliance team blocked their current approach to leveraging the sensitive data in the cloud. They required anonymization in order to leverage the full data sets. In addition, they required full auditing for every use case deployed.

## Immuta Solution

- Self-service data access
- Fine-grained access control over sensitive HR data
- Dynamic masking of sensitive data
- 100% auditability of every query

DEMO

