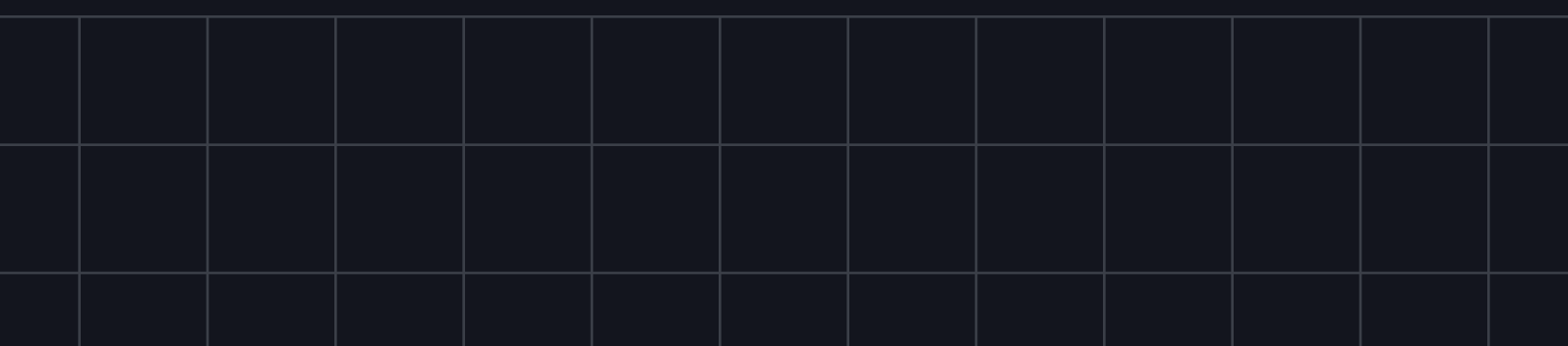




CASE STUDY

# How a Health Data Company Revolutionized Research on Sensitive Data with Immuta



# About the Healthcare Non-profit

The health data non-profit is focused on accelerating research where new and novel data can provide valuable insights and open doors for ethical public problem solving.

The organization was initially established in response to the COVID-19 pandemic and has evolved to tackle other critical policy issues through data. They rely on utilizing commercial, donated, and public-access data sets for analysis by their volunteer corps of data scientists and academic researchers.

With access to this data, different classes of researchers are able to collaborate to produce replicable, modular research that can in turn be published in open-access journals and be used to drive social impact and non-partisan policy work.

## Key Takeaways

- Saved over \$1 million annually in data engineering costs
- Simplified the data request process for researchers, reducing the time to data by 30x, from 90 days to 3 days
- Eliminated the need to manage hundreds of individual data access policies by moving from RBAC to ABAC
- Reduced the number of access control policies to fewer than 10, which can be easily authored by non-technical data governance experts

“As we add more components in a cloud database environment, it’s much cleaner than any sort of on-prem situation we’ve had before. The ability for us to manage access controls, deploy privacy enhancing technologies, and rapidly implement novel frameworks of governance for our research teams has been a breath of fresh air, with no management or overhead costs for adding additional cloud database solutions.”

— Co-Executive Director,  
Healthcare non-profit

## Architecture



## Industry

Healthcare

# Challenge

The organization's mission of enabling ethical public problem solving is predicated upon its ability to provide researchers with new and novel data sets – most of which contain highly sensitive data. For example, under its research program, the non-profit utilized a geolocation data set derived from 50 million Americans' mobile phones during the 2020 Thanksgiving holiday to understand the prevalence of social distancing during the holiday in an effort to control the spread of the virus.

This and other sensitive data, however, can be easily re-identified – an important concern for the non-profit due to legal, ethical, and contractual guidelines that protect against re-identification. It was critical for the organization to find a way to balance data privacy and utility through a mix of data controls, such as masking, and context controls, such as role- and attribute-based access controls.

A key component of the non-profit's approach to research and problem solving is bringing together diverse contributors from interdisciplinary backgrounds. Consequently, they recruit data science and data engineering volunteers from various industries and facilitate their collaboration with academic social science and economics researchers from universities and think tanks.

Given this diverse and ever-changing group of contributors, the non-profit faces the complexity of needing to share sensitive data stored in a com-

partmentalized data lake (using Snowflake as the primary data store) to many different researchers – all of whom need access to different data for varying purposes through the analytics environment. The process involved reviewing each research proposal and then creating new native views and roles in Snowflake to govern access to sensitive data for each research project. This approach quickly became too cumbersome and complicated, particularly in the context of a fast-spreading virus where real-time data was paramount.

With limited resources, it was also critical for the process to be easily managed by the data engineers, who set up the analytics environments for the researchers.

# Solution

To deliver on its mission, the non-profit partnered with Immuta to build an automated data access layer that simplifies data sharing while preserving both data privacy and utility.

When researchers want to begin a new project, they can discover data sets through Immuta's active data catalog and workflows. Self-service workflows enable researchers to request access to data, acknowledge approved usage purposes, request access control changes, and propose new projects.

This approach exponentially simplifies the data request process for their data engineering team. New researchers now have instant access to data, but only the data required to work on their problem scope – reducing time to utilize data from months to hours.

In addition to removing barriers to data access, the non-profit leverages Immuta to automate the enforcement of fine-grained, attribute-based access

controls (ABAC), and privacy enhancing technologies (PETs) on the data stored in Snowflake. Instead of relying on Snowflake controls, nontechnical data governance experts can define robust access policies in Immuta. These policies are then natively enforced when users interact with data in Snowflake.

This new approach greatly simplifies the process of preparing and sharing sensitive data for analytics, eliminating the complexity of creating roles for every researcher and building secure views and complex, role-based functions to implement advanced PETs like format-preserving masking or dynamic k-anonymization.

# Results

By implementing Immuta, The non-profit was able to:

- Save over \$1 million annually in data engineering costs
- Simplify the data request process for researchers, reducing the time to data by 30x, from 90 days to 3 days
- Eliminate the need to manage hundreds of individual data access policies by moving from RBAC to ABAC
- Reduce the number of access control policies to fewer than 10, which can be easily authored by non-technical data governance experts

# About Immuta

Immuta is the **market leader in Data Access**, providing data teams one universal platform to control access to analytical data sets in the cloud. Only Immuta can automate access to data by discovering, protecting, and monitoring data. Data-driven organizations around the world trust Immuta to speed time to data, safely share more data with more users, and mitigate the risk of data leaks and breaches. Founded in 2015, Immuta is headquartered in Boston, MA.

