IMMUTA® | Public Sector

WHITE PAPER

Self-Service Analytics for Data Mesh in the Public Sector

Table of Contents

Introduction

In the past decade, the enactment of both the Data Act and the Evidence-Based Policy Making Act kicked off Federal Agencies' journeys toward data modernization. While many have made progress, there's much more work for the majority of government agencies to do in order to meet the goals outlined by the vision in the President's Management Agenda and the initiatives outlined in the National Defense Strategy. Underpinning these goals is a need to rapidly field capabilities and systems that can enable advanced analytics for direct and indirect mission needs.

Gartner defines self-service analytics as:

> A form of business intelligence (BI) in which line-of-business professionals are enabled and encouraged to perform queries and generate reports on their own, with nominal IT support. Self-service analytics is often characterized by simple-to-use BI tools with basic analytic capabilities and an underlying data model that has been simplified or scaled down for ease of understanding and straightforward data access.

Self-service analytics has been called out in the Federal Data Strategy as an Optimized Activity to occur in the 2026-2028 timeframe. By enabling self-service analytics, organizations are looking to accelerate time to insight with less reliance on IT, and enable data discovery and exploration by business analysts.

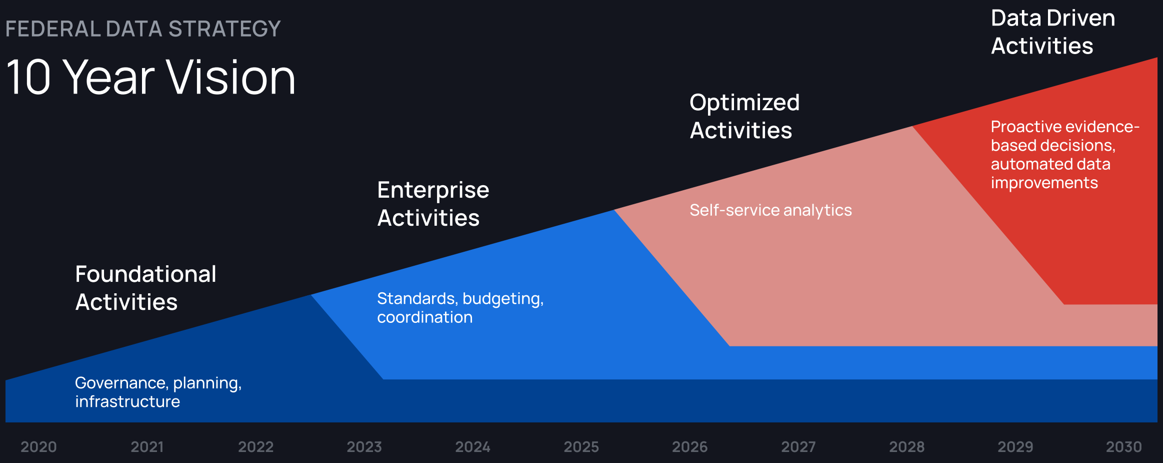


Figure 1: Federal Data Strategy - 10 Year Vision. Source: https://strategy.data.gov/2021/action-plan/

Self-Service Efforts in the Public Sector

Currently, across the public sector, mission leaders are furiously working to support key data-focused mission objectives by enabling users across agencies to operationalize data rapidly for a myriad of analytics use cases. These use cases range from simply improving decision support capabilities through more effective data visualization and dashboards, to enabling data scientists to rapidly build and test artificial intelligence (AI) models that can be deployed into production for national security missions.

There are also efforts across agencies' enterprise IT departments to break down "data silos" through employing modern constructs and technologies like Data Mesh, Data Fabrics, Data Clouds, Lakes, Warehouses and Lakehouses. In any case, the evolving volume, variety, and velocity of data, along with the rapid adoption of modern-stack, cloud-based technologies has created the ability to address these needs in a cost effective and interoperable manner that is expediting data-driven modernization like never before. Ultimately, the time is now for mission leaders to empower users across their organizations to expedite mission outcomes by learning from the innovators that are already delivering results.

There are two efforts simultaneously in play across the Federal Government. First is the establishment of data catalogs as a means to enable data discovery within the enterprise. Second is the establishment of shared-service platforms to handle the storage, management, and processing of data, centralizing analytics tools and capabilities to support a wide range of use cases and user needs. The US Department of Defense's Advana platform and the Air Force's VAULT program are examples of this. While data catalogs allow people to discover data and the platform allows people to perform analytics, the biggest bottleneck to self-service analytics is getting access to the data. In fact, Immuta's 2023 Data Engineering Survey found that 43% believe data access challenges are slowing down analytics processes.

The most innovative agencies have found that while the technology resources are available to process and manage data at scale, a self-service approach must be employed to expedite actual operational impact, eliminate inefficiencies, and empower data consumers. Key to enabling self-service is a data architecture that simplifies access, and balances security and utility through automated access policy enforcement. This is especially critical for Federal agencies that are required to adopt Zero Trust architectures, but is also central for data modernization.

This paper outlines core design and architectural concepts that agencies can leverage to drive mission impact through self-service analytics.

Challenges to Modernizing Data Security and Analytics

To unlock data for self-service analytics, organizations need an automated data access management tool. Today, most organizations rely on in-house or custom solutions when implementing data access control – 44% of those surveyed rely on native access controls, and just 26% have fully automated access control systems.

These controls often either do not provide granular access control, meaning access to data is over-provisioned, or tightly restrict access to just a few users, leading users to create multiple copies of data or write complex SQL statements. Each of these has several pitfalls, including limited ability to share and access data, loss of data integrity due to out-of-date copies, and duplicated or manual policy implementation efforts for every data source technology. Additionally, the manual nature of this work introduces risk that policies will be inconsistently or haphazardly enforced, and that data teams will be unable to keep up with constantly evolving compliance rules and regulations.

For agencies working to modernize, enabling self-service analytics is more complex than just deploying an analytics tool and letting analysts have at it with data. First, data owners need to ensure that the data being analyzed is clean and accurate. There is no escaping "garbage in, garbage out" through technology.

Next, the data needs to be discoverable by users across the enterprise with proper metadata so they can easily find and understand it. Users must be educated in data literacy and empowered to perform analytics. Without a literate user base, insights cannot be derived.

Finally, agencies must establish a data governance program to prevent chaos and ensure data security. Without this final piece, organizations expose themselves to the risk of leaking sensitive data to unauthorized users.

What Can Agencies Do to Secure and Modernize Data Use?

Immuta provides automated data security management that dynamically unlocks access to authoritative data for more users. By integrating with data catalogs such as Collibra, Alation, and Informatica, as well as identity management systems, Immuta is able to leverage user metadata to enforce attribute-based access control (ABAC) policies.

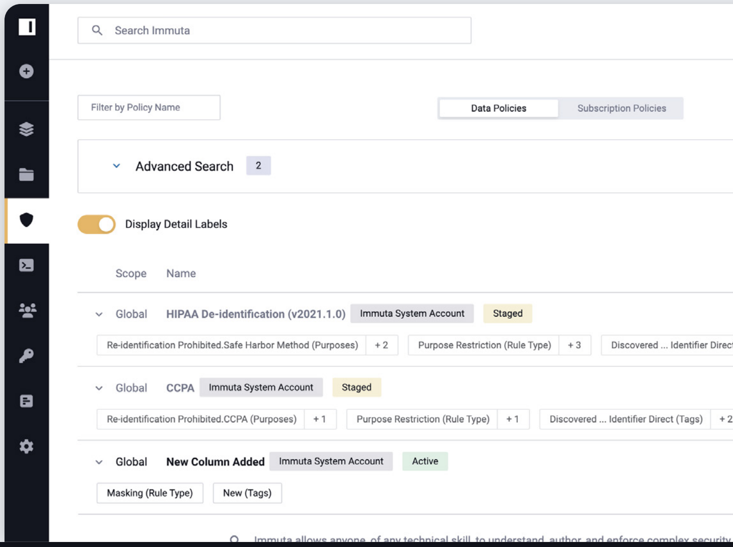
Immuta also automatically discovers and tags over 60 types of sensitive data, such as personally identifiable information (PII), protected health information (PHI), and financial data, and enables organizations to create custom tags to detect other types of sensitive data. Drawing upon user, object, environmental, and action-centric attributes, data stewards are able to create plain language data access and privacy policies in the Immuta UI. Rules may be written to implement data masking at the column- and cell-level, as well as to perform row access filtering, and privacy enhancing technologies (PETs) like k-anonymization and randomized response provide enhanced data protection. These advanced techniques are mathematically proven to preserve data accuracy and privacy during analysis. This is in direct contrast to legacy masking techniques that destroy the underlying data, rendering it useless in analytical use cases. With ABAC, a single Immuta policy can replace more than 100 roles, saving time and reducing security risks.

Adding a further layer of granularity, Immuta's purpose-based access controls (PBAC) ensure specific data sets are accessed only under legal purposes as defined in data use agreements, without requiring copying or moving of data. PBAC thus supports purpose specification and limitation requirements as set forth in data protection and healthcare legislation, as well as department-, agency-, office-, and even team-level data access rules. Purpose statements attached to projects can restrict data access by aligning with a specific data sharing agreement's language. Immuta can enforce who can access what data and under which agreement, require users to accept the language and terms of each agreement, and where applicable, allow users to switch between projects they have accepted and are authorized to access. The ability to meet purpose specification requirements is central to meeting other data protection requirements, such as data minimization, and ultimately mitigating re-identification risks using PETs.

The key to Immuta policies is that they follow a write-once and enforce-everywhere approach. This means that when a policy is written in Immuta, a single policy is enforced on all integrated data platforms without the need for database administrators (DBA) or duplicate policies for each data platform. Policies are also automatically added to new data sources that are integrated into Immuta and when new columns are added to existing data sources. This automation ensures data is always consistently protected across technologies and as data changes over time. As a result, Immuta reduces the risk of unauthorized exposure to sensitive data stemming from the inability to make real-time changes.

Immuta's Data Security Platform gets the right data into the right hands faster, enabling data teams to break through data access barriers and achieve both scalability and security.

Ready to get started?
Request a demo today.



Search Immuta
Filter by Policy Name
Data Policies
Subscription Policies
Advanced Search
2
Display Detail Labels
Scope
Name
Global
HIPAA De-identification (v2021.1.0)
Immuta System Account
Staged
Re-identification Prohibited.Safe Harbor Method (Purposes)
+ 2
Purpose Restriction (Rule Type)
+ 3
Discovered ... Identifier Direc
Global
CCPA
Immuta System Account
Staged
Re-identification Prohibited.CCPA (Purposes)
+ 1
Purpose Restriction (Rule Type)
+ 1
Discovered ... Identifier Direct (Tags)
+ 2
Global
New Column Added
Immuta System Account
Active
Masking (Rule Type)
New (Tags)
Immuta allows anyone, of any technical skill, to understand, author, and enforce complex security

NAICS CODES
541511, 541330,
541512, 541513,
541519, 541990

CONTRACT VEHICLES
GSA Schedule 70,
NASA SEWP, FirstSource,
ITES SW2

publicsectorsales@immuta.com

IMMUTA®

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

010623