

# Collibra Data Intelligence Cloud for Federal Agencies

## Overcoming data challenges

Federal agencies operate in a highly complex and regulated data landscape with petabytes of data residing in hundreds of disparate data sources. These sources are spread across on-premises, hybrid, and even multi-cloud environments.

Data citizens – including Chief Data Officers, data scientists, data stewards, and government employees at large – are unable to easily find or understand the data they need. It is difficult for users to ascertain if the data comes from a trustworthy source, whether it is certified, or approved for use. In other words, data citizens are unable to fully appreciate the corpus of data and its relevance for use in their applications.

For any agency, the inability to find the right data combined with data integrity issues, a lack of data governance, transparency, and traceability of data can adversely impact their ability to serve constituents or critical missions in a timely, efficient, and secure manner.

To address these issues, the Federal Data Strategy and the Foundations for Evidence-based Policymaking Act provide a framework for agencies to overcome data silos. The goals of these initiatives are to make data available, accessible, re-useable, and shareable. This requires the implementation of a rigorous data intelligence and management strategy that accounts for all data assets created, collected, controlled, and maintained by each agency.

**Collibra Data Intelligence Cloud is an intelligent data platform that leverages the power of machine learning and provides a holistic and integrated approach to cataloging, governing, protecting, managing and collaborating with data at scale across on-premises, hybrid and multi-cloud environments.**

## Drive innovation with trusted data

With Collibra Data Intelligence Cloud, federal agencies can build a common foundation for data. It helps speed data discovery, improve data literacy (understanding of data), and drive greater collaboration with trustworthy data both within and across federal agencies.

Collibra Data Intelligence Cloud is built on a zero-trust architecture allowing agencies to securely and quickly govern data, protect classified and sensitive information, and enable data integrity at scale. Security and data protection are key tenets of Collibra Data Intelligence Cloud. As a commitment to serving federal agencies requirements, Collibra is also FedRAMP Moderate authorized.

Collibra Data Intelligence Cloud serves the most relevant data to key stakeholders to assist them in making critical decisions across the civilian, military, and intelligence communities. Users in these communities are accelerating data-driven decisions, enabling mission-critical applications, mitigating operational risks, operationalizing government mandates, policies, and laws including:



**Federal Data Strategy**



**Self-Service analytics and AI governance**



**Evidence-based Policymaking Act**



**Financial and regulatory reporting**



**Open Government Data Act**



**Digital transformation and data modernization**

# Enable strategic use cases



## Accelerate data democratization and self-service analytics

- Empower data citizens to rapidly discover, understand and collaborate with fully governed data that's primed for self-service analytics and AI use.
- Enable a federated data marketplace built on a modern data mesh architecture for faster time-to-insights.
- Leverage popular business intelligence (BI) tools such as Tableau in conjunction with Collibra Data Catalog and enable data citizens to obtain a more complete view of data with granular details for more consistent, accurate and transparent reporting.
- Make better and more informed decisions faster leveraging high-quality data to support critical missions and ultimately better serve constituents in a timely and secure manner.



Figure 1: Easily collaborate with data leveraging crowdsourcing capabilities



## Enable explainable AI and build models you can trust

- Create a trusted repository of metadata (data about data) that includes AI model techniques, data inputs, data features, expected outputs, and more.
- Empower data scientists to easily discover trustworthy data with rich business and technical context so they can build robust data pipelines and train models.
- Trace lineage of datasets and models with detailed information to better understand data transformations throughout the data lifecycle.
- Leverage machine learning (ML)-enabled data quality and observability to help detect data drift, outliers and patterns to ensure accuracy and performance of AI and analytics models over time.



## Improve transparency and data integrity at scale

- Implement systematic data governance, quality and privacy controls that remove human error and mitigate data-related, operational and regulatory non-compliance risks.
- Proactively detect anomalies and inaccuracies in data across agency departments and systems including HR, finance, and more.
- Mitigate compliance and cybersecurity risks by applying data policies and controls consistently across all datasets.
- Easily locate, classify and protect sensitive data including Personally Identifiable Information (PII) and Protected Health Information (PHI) with role-based access control.

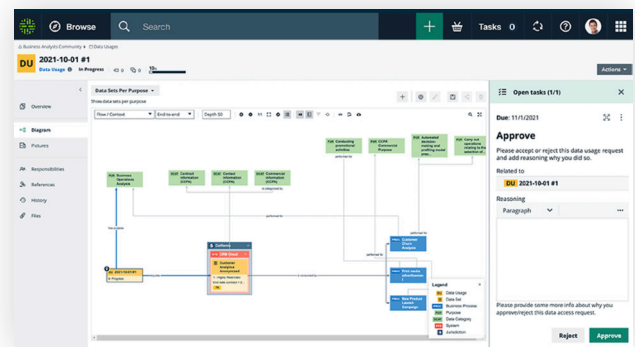


Figure 2: Visualize and understand data dependencies from source to target with end-to-end data lineage



## Speed digital transformation and data modernization

- Quickly determine what impact migrating a particular dataset to a data lake or modern data warehouse may have on upstream and downstream systems, as well as prioritize migration of datasets in the right sequence and securely.
- Simplify data onboarding to maximize productivity by cataloging and managing the data in your data lake and data warehouse from a single centralized location.
- Maximize the impact of your data migration by setting up controls to ensure that only clean data that meets your data quality standards is migrated.
- Obtain end-to-end visibility into your data residing in data lakes and data warehouses in AWS GovCloud.



# The Collibra Data Intelligence Cloud Advantage



Federal agencies as well as the intelligence community relies on Collibra Data Intelligence Cloud to gain in-depth visibility into their data ecosystem across on-premises, hybrid and AWS GovCloud without moving data, so that data assets stay protected.

## Govern. Trust. Access.

### Data Catalog

- Catalog
- ML curation
- Ratings / reviews
- Data marketplace

### Data Lineage

- Business lineage
- Technical lineage
- Impact analysis

### Data Governance

- Business glossary
- Stewardship
- Data dictionary
- Reference data
- Data helpdesk

### Data Quality

- Continuous monitoring
- Adaptive rules

### Data Privacy

- GDPR
- CCPA

## Active metadata graph

### Core services



Flexible operating model



Native connectors



Reports and dashboards



Collibra everywhere



Workflow engine



Collaboration



APIs



Collibra Insights

## Core capabilities



### Rapid data discovery and self-service data access

- The intelligent data catalog is purpose-built to empower data citizens, so they can easily find and understand the data they need with rich business and technical context. With built-in crowdsourcing capabilities, users can easily collaborate on data. They can rate, comment on, and prioritize data sets for use as well as provide feedback.
- A powerful contextual search engine and easy-to-use, intuitive data shopping experience enables data citizens across the agency to easily locate and shop for fully-governed and trustworthy data they need.



### End-to-end data lineage and visibility

- Automated data lineage enables data citizens to dynamically visualize the flow of data from source to target as well as understand all the data dependencies at a granular level.
- For instance, data citizens can better understand financial disclosures and trace lineage of each data set with rigorous audit trails for reporting purposes. An active metadata graph – that connects business, technical and privacy metadata with data quality and lineage - provides rich content and context to users.



### Predictive data quality and data integrity at scale

- Machine learning (ML)-powered data quality and observability enables proactive detection of anomalies in data such as missing records, values as well as broken relationships across tables or systems resulting in rapid resolution.
- If data quality drops below a set threshold, Collibra automatically creates a data quality issue and notifies the right people to handle the issue according to the desired priority level.



### Comprehensive and adaptive data governance

- Collibra data governance is a one-stop-shop for data, regulatory compliance, AI and analytics governance. Built-in data governance enables users to develop a consistent data taxonomy that includes regulatory specific rules, physical and/or logical data dictionary, glossary of terms, out-of-the box data domains, & more.
- Pre-built templates, reference data management, stewardship management and intuitive workflows enable automation of governance processes and provide a framework for cross-functional teams to establish a common understanding of data to ensure consistency.



### Embedded privacy by design and data protection

- With Collibra, federal agencies can enforce comprehensive data privacy and protection policies in adherence with regulatory stipulations and corporate best practices for data retention and usage.
- ML-powered automation enables rapid classification of data by category, sensitivity and user type at scale. Advanced capabilities in data privacy and protection will help ensure sensitive data is easily identified and protected with role-based access control.

**Collibra is the Data Intelligence company**

For more information, visit us at [collibra.com](https://collibra.com)

We accelerate trusted business outcomes by connecting the right data, insights and algorithms to all Data Citizens. Our cloud-based platform connects IT and the business to build a data-driven culture for the digital enterprise. Global organizations choose Collibra to unlock the value of their data and turn it into a strategic, competitive asset. We have a diverse global footprint, with offices in the US, Belgium, Australia, France, UK, Czech Republic and Poland.