

Secure and unlock unstructured content for greater mission success

The challenge: Vast content repositories

Federal agencies possess a staggering amount of file content, also called unstructured data, which is estimated to make up 80%-90% of their total data volume. For data governance teams, effectively managing, securing and utilizing these vast data stores presents a significant challenge, as unstructured data is inherently difficult to analyze, manage and protect.

The solution: Collibra Public Sector and Ohalo integration

Collibra Public Sector and Ohalo deliver a comprehensive solution designed to manage structured, semi-structured and unstructured data. This partnership helps federal agencies unlock the power of all of their data to better serve citizen and warfighter missions.

Ohalo Data X-Ray transforms unstructured data into structured insights and integrates seamlessly with Collibra, enhancing governance through automated classification, metadata tagging and risk assessment. By automating the sensitive data tagging and labeling of unstructured data, Data X-Ray enriches and expands the view of data within Collibra. This allows federal agencies to discover, classify and govern their unstructured data at scale, enabling more compliant and secure data and AI applications.



Key benefits and capabilities

Automated data discovery, classification and tagging

- Tags and labels sensitive and mission-critical data across enterprise storage, providing better visibility and control
- Detects the “Alphabet Soup” of regulated data, including CUI and classified information, sensitive information and financial data, to ensure compliance and security
- Aligns unstructured data with business glossary terms for seamless governance

AI and data science enablement

- Curates files to provide context for AI Retrieval-Augmented Generation (RAG)
- Monitors the quality of unstructured data
- Extracts and makes key content attributes searchable using AI
- Maintains audit-ready governance records inside Collibra for all AI and data workflows

Data privacy and records compliance automation

- Keeps an up-to-date index of people, organizations and sensitive entities to support privacy governance
- Automates responses for Data Subject Access Requests (DSAR) and Freedom of Information Act (FOIA) requests with entity-based search
- Syncs Collibra asset tagging with Data X-Ray to streamline records and data retention use cases

Deployment in mission-critical and air-gapped environments

- Can be deployed wherever your data resides
- Ensures no data ever leaves your environment
- Allows non-AI functions to run independently in degraded environments that lack GPU or AI API access

Addressing key federal use cases

With the combined power of Collibra Public Sector and Ohalo, federal agencies can address their most pressing use cases.

Objective	Capabilities
Govern and enable AI	Curate, classify and enrich data for AI; enhance data quality and provide access governance for RAG, AI agents and assistants
Comply with security policies	Automatically identify PII, IP and classified data; search and extract data from unstructured sources to apply data retention policies
Find and discover data	Locate specific data for DSARS, FOIA requests and legal discovery needs
Implement zero trust	Discover, label, and tag mission-critical assets and integrate with security and IAM stacks to enable an Attribute-Based Access Control (ABAC) framework

 Collibra Public Sector

Collibra frees your data from the constraints of silos by unifying data and AI governance across every system and bringing business and technical users into the fold. It gives you a higher degree of compliance paired with more autonomy, so your users can trust, comply and consume.

 OHALO

Built for multi-cloud and hybrid environments, Data X-Ray integrates into complex agency infrastructures. Ohalo’s team has deep expertise in handling governance challenges, ensuring structured control over complex data ecosystems.