Oracle Cloud Infrastructure Data Catalog

Oracle Cloud Infrastructure Data Catalog helps customers gain greater insights from their data in Oracle Cloud and beyond. It enables data professionals such as analysts, data scientists, and data stewards to discover, organize, assess, and enrich data in a truly self-service and governed way. This minimizes time spent searching for data and maximizes time spent extracting its value for analytics and data science projects.

Included within every Oracle Cloud Infrastructure subscription at no extra cost, Data Catalog is a metadata management solution that creates an organized, searchable inventory of data assets based on technical, business, and operational metadata. It also supports data governance by helping customers find, understand, and track their cloud data assets.

Key Features

- Technical metadata harvesting for a variety of enterprise systems
- Metadata enrichment with free form tags and business terms for a holistic view
- Searchable, standardized inventory of data assets
- Business glossary to provide common meaning across disparate datasets and users
- Oracle Cloud Infrastructure Data Catalog APIs and SDKs

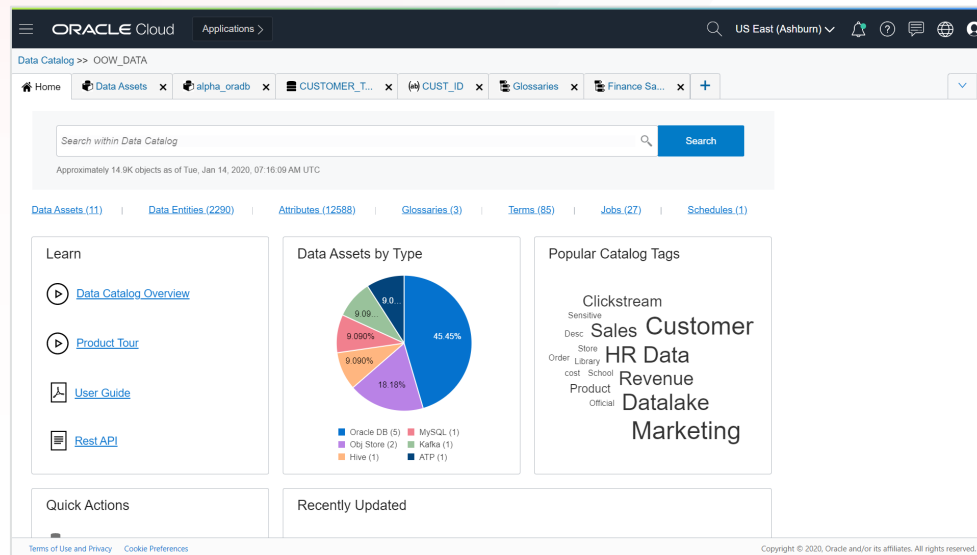## THE CHALLENGE OF GOVERNING ENTERPRISE DATA AS AN ASSET

In the new digital economy, the convergence of cloud, big data analytics, and artificial intelligence/machine learning is driving new opportunities, all with data at the core. Organizations must work to derive value from data to gain competitive advantage and operational excellence. But there are challenges.

As data volumes explode, finding trustworthy data is still a manual and time-consuming process.

Organizations are also migrating from on premises to cloud to multi-cloud. In such a complex journey, it is critical to understand the data landscape: the scope, the source, and the impact. With increasing rules and regulations around privacy and usage, better data governance practices are a must.

Oracle Cloud Infrastructure Data Catalog is a single collaborative solution for data professionals to find, understand, and track trusted data in Oracle Cloud and beyond. It allows users to collaborate, enrich, and manage the enterprise view of data assets by capturing their subject matter expertise for accuracy on key elements (e.g., business meaning, context, usefulness, quality levels, fitness for use, origins, and policy constraints).
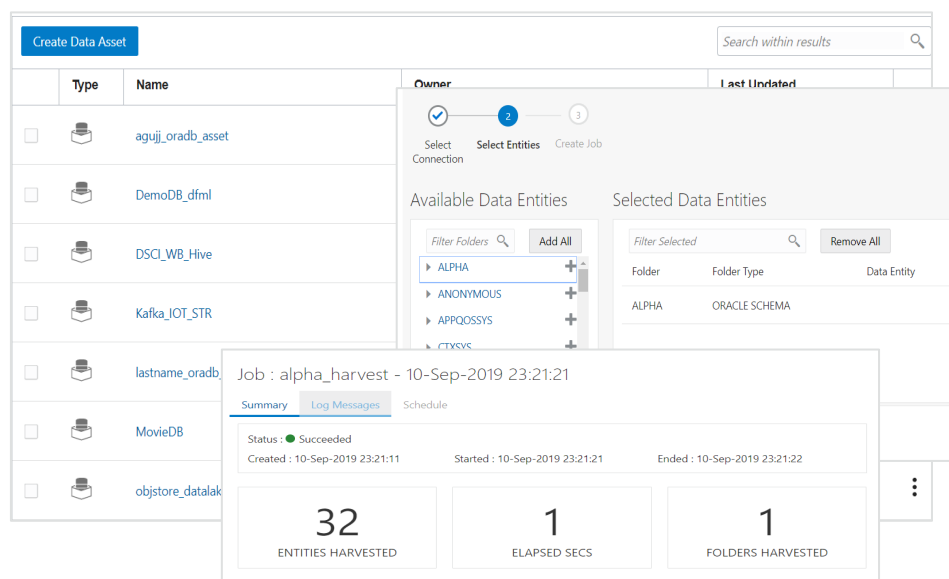

Key Business Benefits

- Gain better visibility into data assets in the enterprise to establish trust and transparency

- Quickly find and explore relevant data for analytics and data science projects

- Capture business vocabulary and context for data assets for better search and discovery and a holistic view of data

- Use APIs to integrate capabilities into other Oracle Cloud Infrastructure services or external applications

HARVEST METADATA FROM A VARIETY OF ENTERPRISE CLOUD SYSTEMS

Oracle Cloud Infrastructure Data Catalog harvests technical metadata information from connected data assets such as Oracle Cloud Infrastructure Object Storage, Oracle Autonomous Data Warehouse, Oracle Autonomous Transaction Processing, Oracle Database, MySQL, Hive, and Kafka. It then gathers details about available data entities and attributes into the catalog and creates a searchable inventory.
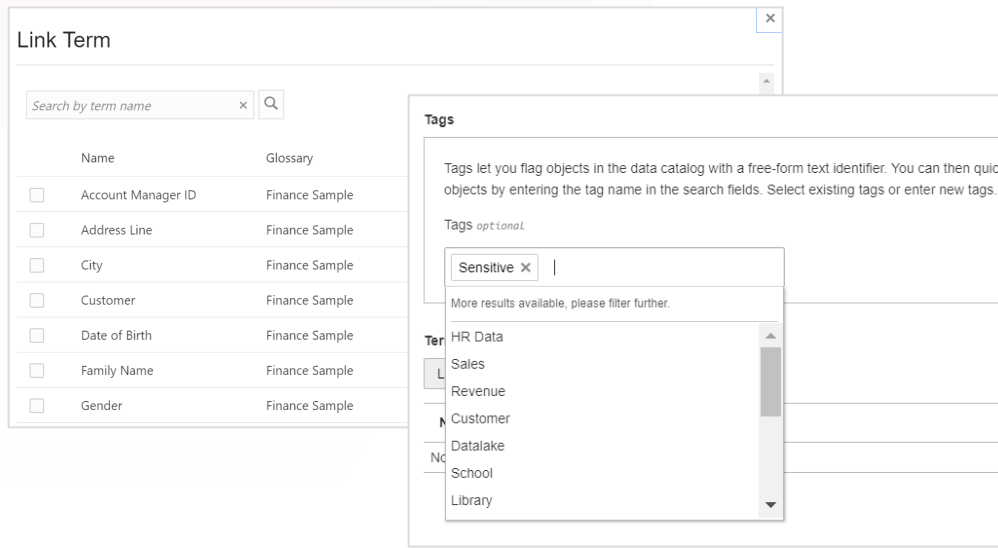
METADATA ENRICHMENT

Different users and subject matter experts can collaboratively enrich technical information with business context to capture and share their knowledge. Data entities and attributes can be tagged or linked to business terms to capture tribal knowledge and provide a holistic view. These enrichments also help with classification, search, and data discovery.
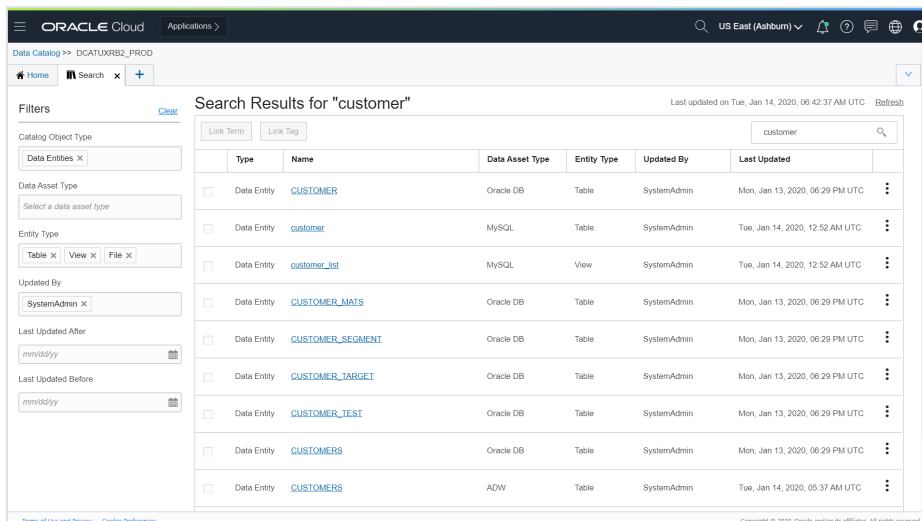

SEARCHABLE DATA ASSET INVENTORY

Oracle Cloud Infrastructure Data Catalog creates a powerful, searchable, standardized inventory of the available data sources, entities, and attributes. For searching, users can enter technical information, user-defined tags, or business terms to search. Flexible searching and filtering options allow users to quickly find relevant sets of data for data science, analytics, or data engineering. Users can also browse metadata based on technical hierarchy of data assets, entities, and attributes.

Related Products

- Oracle Cloud Infrastructure Object Storage
- Oracle Autonomous Data Warehouse
- Oracle Autonomous Transaction Processing

BUSINESS GLOSSARIES

One of the first steps towards effective data governance is establishing a common understanding of business concepts across the organization and their relationships to the data assets within the organization. Oracle Cloud Infrastructure Data Catalog includes capabilities to collaboratively define business terms in rich text form, categorize them appropriately, and build a hierarchy to organize this vocabulary. Users can create parent-child relationships between various terms to build a taxonomy. They can also set business term owners and approval status so that users know who can answer their questions regarding the terms. Once created, users can then link these terms to technical assets to provide business meaning and use them for searching as well.

DATA CATALOG API AND SDK

Many of the Oracle Cloud Infrastructure Data Catalog capabilities are also available as public REST APIs to enable integrations such as:

- Searching and displaying results in applications that use the data assets

- Looking up definitions of defined business terms in the business glossary and displaying sthem in reporting applications

- Invoking job execution to harvest metadata as needed

CONCLUSION

Oracle Cloud Infrastructure Data Catalog provides a single collaborative solution for data professionals to collect, organize, find, access, enrich, and activate technical, business, and operational metadata to support self-service data discovery and governance for data assets in Oracle Cloud. How does your enterprise understand and leverage its data? How do you better support self-service analytics without compromising governance requirements? Try Oracle Cloud Infrastructure Data Catalog, available within every Oracle Cloud Infrastructure subscription, to start discovering the value of your data today.

CONNECT WITH US

Call +1.800.ORACLE1 or visit oracle.com.
Outside North America, find your local office at oracle.com/contact.

B blogs.oracle.com/oracle   **f** facebook.com/oracle   **y** twitter.com/oracle

Integrated Cloud Applications & Platform Services

Oracle is committed to developing practices and products that help protect the environment