

PRODUCT OVERVIEW

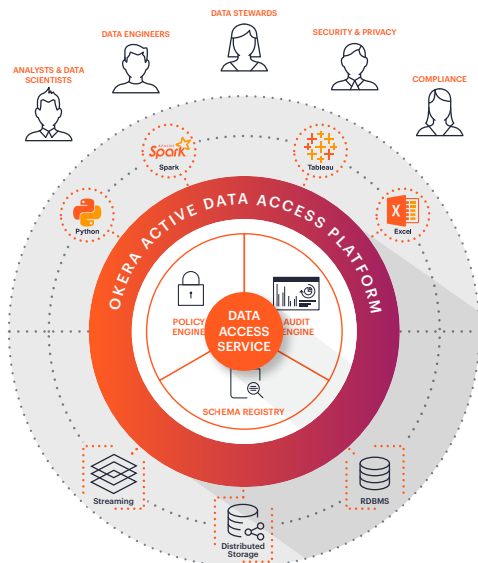
Okera Active Data Access Platform for Amazon Web Services (AWS)

Okera solves one of the biggest challenges with heterogeneous, multi-vendor data platforms at scale on AWS: opening data access for innovation while ensuring proper governance and compliance. Without the ability to scale governance, enterprises are limited in what they are able to do with new data platform technologies or run the risk of their security and governance controls falling short.

The Okera Active Data Access Platform unifies and manages access for data consumers across a multi-datastore and multi-tool environment. It reduces friction between agility and governance and the people involved-data consumers, owners, and stewards.

Product Overview

The Okera Active Data Access Platform consists of the Data Access Service and Catalog Services. The Catalog Services include Automated Schema Registry, Policy Engine and Audit Engine. The platform is modular in nature and exposes standard APIs making it easy for data platform teams to integrate it within their environment and also gives them the choice to use other services as they see fit.



Key Benefits

FLEXIBILITY

Work with any analytic tool and engine including native AWS and partner services.

PROTECTION

Ensure granular access control with data obfuscation and high-performance at scale.

VISIBILITY

Gain usage insight into access patterns with auditing.

Okera Serves All Data Users

CONSUMERS

Data consumers get fast and secure access to data without any change in the analytic tool they are most comfortable with.

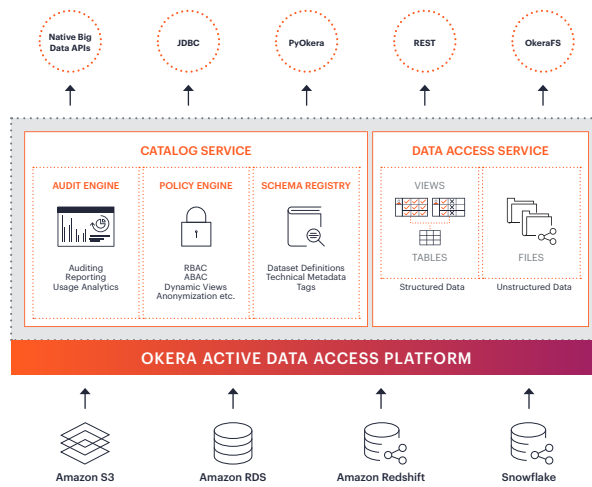
OWNERS

Data owners get an easy way to publish datasets and thereafter get visibility into how the data is being used.

STEWARDS

Data Stewards define and enforce access policies centrally at a fine-grained level, understand access levels, as well as audit historical data access activity. There is no more manual hassle to deal with unwieldy and complex JSON-based IAM policies.

Democratize analytics while enforcing governance



DATA ACCESS SERVICE

The Okera Data Access Service is a scalable, fault-tolerant distributed service, and your analytics tools interact with it for both structured and unstructured data. The Data Access Service does the I/O and provisions data to the analytics tools after applying schema, fine-grained access policies and other transformations (UDFs, pseudonymization, masking, etc.) with high performance. Data is provisioned in an easily consumable form. For structured data, it has the familiar abstraction of tables and views. Unstructured data is consumed in the form of files formats the user may request. Different retrieval, streaming, and analytics tools like Spark, Python, SQL engines, Notebooks, and business intelligence tools like Tableau, Microsoft Power BI, and Microsoft Excel interact with this service.

AUTOMATED SCHEMA REGISTRY

The Okera Automated Schema Registry provides the ability to automatically discover, store, and query technical and operational metadata on datasets available to data consumers. Schemas, dataset sizes, who owns them, tags, annotations, basic quality metrics are some of the information that the schema registry contains.

Users and applications can access the registry via the Hive Metastore API, REST APIs or a GUI. It typically becomes the central schema registry that is shared across multiple analytics tools.

POLICY ENGINE

The Okera Policy Engine provides the ability to define and manage data access policies that can be applied on the fly as data is accessed. These policies can be defined at several granularities: database, dataset, rows, columns and even cells.

Role-Based Access Control (RBAC)

Permissions are based on roles, or personas, who need to perform specific data-centric tasks. RBAC is typically combined with IAM systems, such as LDAP-based directories, to tie user groups into the role-based permissions.

Data Obfuscation for Differential Privacy

Sensitive data can be dynamically protected using obfuscation functions including anonymization, pseudonymization, redaction, and masking.

Attribute-Based Access Control (ABAC)

ABAC makes it easy to assign access permission with tags to enrich data sets with business context. You can then automate business metadata tagging and policy enforcement to assign policy at scale.

AUDIT ENGINE

The Okera Audit Engine gives a detailed view of what's happening with your data. This includes but is not limited to information about different user activity, most popular datasets and most commonly used tools.

Automating Data Privacy

The capabilities of Okera combine to ensure automated data privacy:

AUTOMATED DATA DISCOVERY

The Automated Schema Registry crawls the data store and automatically discovers metadata. The Policy Engine tags data based on customizable patterns with labels that were defined for ABAC.

AUTOMATED POLICY ENFORCEMENT

After sensitive datasets are tagged through the Policy Engine with ABAC, data access and obfuscation is applied on the fly and at scale to ensure Differential Privacy.

Regulatory Compliance: GDPR and CCPA

Okera provides you the necessary controls to not only adhere to regulatory compliance, but also to use those standards as foundational abilities to drive new business initiatives.

ABOUT OKERA

Okera enables the management of data access and governance at scale for today's modern cloud data lakes. Built on the belief that companies can do more with their data, Okera's Active Data Access Platform (ODAP) enables scalable fine-grained data protection and visibility on data lakes for both structured and unstructured data. This allows agility and governance to co-exist and gives data consumers, owners and stewards the confidence to unlock the power of their data for innovation and growth. Enterprise organizations receive immediate value from Okera which can be implemented and deployed in less than a day. Okera is headquartered in San Francisco and is backed by Bessemer Venture Partners, Felicis Ventures, and Capital One Growth Ventures.

Learn more at www.okera.com or contact us at info@okera.com.