

## DATA LAKE GOVERNANCE AND SECURITY

# Protect Your Data Lake With Fine-grained Access Control and Governance

---

### Big Data Creates Big Challenges

“Big Data” is getting bigger by the second. Organizations are now routinely collecting, storing, processing and analyzing terabytes and petabytes of data in data lakes. Data scientists, analysts and others are seeking to leverage these huge data stores to gain every possible competitive advantage. Data lakes—storage repositories that hold vast amounts of raw data in its native formats, to be held until it is processed—are changing enterprise data management for the better because the data does not need to be restructured to be stored, and the separation of storage and compute opens up several possibilities multiple types of workloads that can be run on the same data.

### Importance of Access Management and Governance as Data Lakes Mature

It’s easy to understand the enthusiastic embrace of data lakes by data-driven organizations. However, early adopters tend to focus first on the new capabilities provided by the technologies, leaving other enterprise requirements, such as access management and governance, to be addressed later.

It’s typical to see enterprises start adoption with a few users and use cases and get comfortable with this paradigm. As adoption and data grow, users and workloads are onboarded and security and governance of data become a priority. This tends to slow down adoption and hurt user productivity as governance trumps agility. According to IDG’s 2015 Big Data and Analytics Survey, “As organizations collect, store, and analyze increasing amounts of data from new and existing sources, security becomes of greater concern. Nearly 35% of survey respondents either aren’t sure or don’t think that their existing security solutions and products provide adequate data security.”

---

### Data Lakes Benefits

Data Lakes separate storage and compute to deliver numerous benefits to organizations:

#### FLEXIBILITY AT MULTIPLE LEVELS

- Data type: Store and access structured and unstructured data assets in native formats including CSV formats, file storage, relational databases, rich media (e.g. audio/video), time-series log data, and more
- Workload type: Support multiple types of workloads

#### SCALABILITY

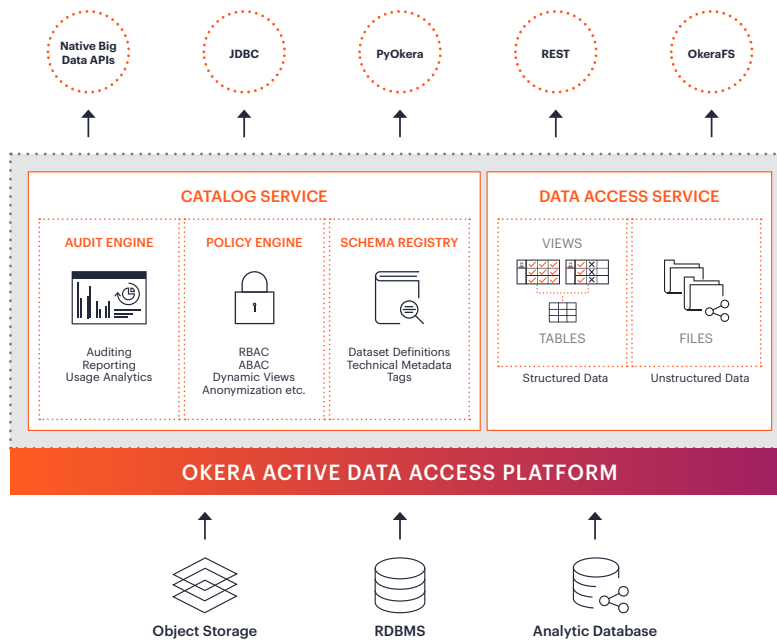
Scale to multiple petabytes and beyond

#### COST EFFECTIVENESS

Leverage several technologies at scale with a much lower price point than ever before

#### OKERA PROTECTS DATA LAKES BY

- Simplifying and unifying access to data from different tools
  - Enabling access control at different granularities: rows, column and even cells
  - Dynamic anonymization to help you meet compliance requirements without inhibiting access
  - Enabling governance and auditability of your data at scale
  - Lowering the risks of data propagation by reducing the spread of replicas
  - Ensuring regulatory compliance: GDPR, CCPA, and so on
-



## Okera Protects Access to Data Lakes

Security and governance are essential to every successful data lake implementation at scale. Data platform teams have to service the enterprise governance requirements in order to scale the lakes and make them mainstream. The downside risks and costs of failing to protect access and data are high in terms of regulatory penalties, reputational damage and even loss of customers. Okera enables you to address these requirements at any point during your data lake evolution.

## Okera Active Data Access Platform Capabilities

The Okera Active Data Access Platform provides powerful access and governance capabilities. It easily and rapidly integrates with data lake implementations, enabling your organization to secure all data held in and accessed from your data lake. Specifically, it enables:

### EASY ACCESS TO DATA

While flexibility is a key benefit offered by data lakes, it is also a potential source of vulnerability due to the diversity of data held in the lake. Okera's Active Data Access Platform provisions data to users with an easy-to-use abstraction of datasets. Users can consume these datasets using their preferred tools without having to deal with the underlying physical systems, data layouts and formats.

### FINE-GRAINED ACCESS CONTROL

By implementing the Okera Active Data Access Platform, customers, data producers, and data stewards can easily request, grant, and manage access rights at different granularities: rows, columns and even individual cells. The platform also enables dynamic data obfuscation, including anonymization, pseudonymization, redaction, and masking of data on the fly. Additionally, attribute-based access control (ABAC), makes it easy to assign access permission

with tags to enrich data sets with business context. ABAC may be used to automate business metadata tagging and policy enforcement to instantly understand business context of your data at scale, and automate securing personally identifiable information.

### COMPREHENSIVE GOVERNANCE

For any big data technology solution to be considered secure, you must be able to monitor data usage by each individual, and to produce a complete audit trail at scale. Okera's Active Data Access Platform enables you to do this in addition to getting a clear view of who has access to what data at any point in time.

*“Data and analytics leaders need to take an aggressive approach that creates an appropriate balance between data collection and data connection.”*

—ROXANE EDJLALI AND TED FRIEDMAN, GARTNER ANALYSTS

### 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](http://www.okera.com) or contact us at [info@okera.com](mailto:info@okera.com).

© Okera, Inc. 2019 All Rights Reserved. UB-DataLake - MAY2019