DGOV-221

Controlling with Cloudera Data Governance

Cloudera Shared Data Experience and Data Catalog powered by Apache Ranger and Atlas

Course Overview

Course Type
Instructor-Led Training

Level
Intermediate

Duration
2 days

Platform
Cloudera on premises

Topics Covered

- Auditing
- Lineage
- Access Control

About This Training

This course helps customers use Cloudera to address data governance tasks, motivated by the need for compliance with regulations such as the European Union's General Data Protection Regulation (GDPR) and the United States' Health Insurance Portability and Accountability Act (HIPAA).

What You Will Learn

Through instructor-led discussion, demonstrations, and hands-on exercises, you will learn how to:

- Identify which tools in Cloudera to use for key data governance activities
- Organize data objects using classifications and business glossary terms
- Find access history for data objects and policies
- Use Data Catalog Profilers in Cloudera to assist in organizing data objects
- Use Data Catalog to foster collaboration with colleagues
- View and interpret a data object's lineage
- Create and apply resource- and tag-based access control policies
- Create policies for data masking and row-level filtering

Who Should Take This Course?

This course is best suited for data stewards and others who are responsible for, or have an interest in, implementing regulatory compliance or performing typical data governance activities using the Cloudera platform. Familiarity with basic data governance concepts is helpful, but not required.

DGOV-221

Controlling with Cloudera Data Governance

Introduction

- Team activity: Team Introductions

Data Governance Overview

- What Is Data Governance?
- Basic Concepts
- SDX: Data Governance in Cloudera

Organizing Data Objects

- Searching for Objects by Type
- Classifications
- Glossary Terms

Auditing

- Auditing Overview
- Viewing Audit Information

Working with Data Catalog

- Data Catalog Overview
- Sensitive Data Profiler
- Defining and Monitoring Data Quality
- Preparing for Audits Using Data Catalog
- Collaborating

Lineage

- Inspecting Lineage
- Propagation and Lineage in Atlas
- Inspecting Lineage in Atlas

Access Controls

- Apache Ranger Basics
- Creating Users and Roles
- Resource-Based Policies
- Tag-Based Policies
- Securing Metadata Objects
- Providing Partial Access

Managing the Data Lifecycle

- Governing the Data Lifecycle