

DGOV-223



End-to-End Data Governance with Cloudera Data Lineage

Course Overview

Course Type

Instructor-led training

Level

Intermediate

Duration

3 day

Platform

Cloudera On premises
Cloudera Data Lineage

Topics Covered

- Introduction to Cloudera Data Lineage
- Auditing and Access Control
- Connectivity and Data Source Integration
- Cloudera Connector Factory.
- Platform Navigation and Data Discovery
- Deep Dive into Lineage and Impact Analysis
- Knowledge Hub and Data Catalog Management
- AI-Driven Insights and Optimization(Cloudera Lineage AI Assistant)

About This Training

This course provides an introduction to implementing data governance and lineage using Cloudera Data Governance and Cloudera Data Lineage. Participants will learn how to manage, control, and trace data across the enterprise while supporting compliance requirements such as GDPR and HIPAA.

The training covers data classification, protection, metadata management, lineage tracking, and impact analysis to improve transparency, regulatory compliance, and data trust.

What Skills You Will Gain

Participants Will Develop the Ability To:

- Understand the role of Cloudera Data Lineage within a modern data governance framework
- Navigate the Cloudera governance and lineage interfaces with confidence
- Connect and analyze data sources for lineage and metadata visibility
- Perform end-to-end lineage analysis for impact assessment and troubleshooting
- Organize and catalog data using classifications and business glossary terms
- Use profiling capabilities to better understand and manage data assets
- Create and apply resource- and tag-based access control policies.
- Implement data masking and row-level filtering policies for compliance
- Leverage metadata intelligence and AI-driven insights for governance optimization

Who Should Take This Course?

This course is designed for data professionals responsible for managing, governing, and understanding enterprise data using the Cloudera platform and Cloudera Data Lineage. It is ideal for data stewards, data governance professionals, data analysts, data engineers, BI developers, and solution architects who are involved in regulatory compliance, metadata management, lineage analysis, and data access control.

DGOV-223

End-to-End Data Governance with Cloudera Data Lineage

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

DGOV-223

End-to-End Data Governance with Cloudera Data Lineage

Introduction to Cloudera Data Lineage & Core Concepts

- What is Cloudera Data Lineage?
- Understanding Data Lineage and its Importance
- Overview of Cloudera data lineage's Key "Spaces"
- Lab: Access to Cloudera Data Lineage Environment

Cloudera Data lineage Connectivity & Data Source Integration

- Overview of Cloudera Data lineage Connectivity
- Cloudera Data Lineage Agent Installation
- Types of Connectors
- Demo: Cloudera Data Lineage Agent Installation
- Demo: Cloudera Data Lineage Metadata Extraction

Cloudera Connector Factory

- Overview of Cloudera Connector Factory
- Cloudera Connector Factory Instructions
- Types of Templates
- Log Troubleshooting & Extraction Validation
- Demo: Cloudera Connector Factory

Navigating the Cloudera Data Lineage Platform & Discovery Space

- Cloudera Data Lineage User Interface Orientation
- Introduction to the Automated Discovery Space
- Lab: Navigating to Cloudera Data Lineage User Interface h Automated Discovery Space

Deep Dive into Lineage Space

- Understanding Data Lineage Types in Cloudera Data Lineage
- Cross system Lineage in Cloudera Data Lineage
- Inner System Lineage in Cloudera Data Lineage
- End-to-End Column Lineage for Impact and Root Cause Analysis
- Lab: Navigating to Cloudera Data Lineage Lineage Space

Live Lineage & Cloudera Lineage AI Assistant

- Introduction to Live Lineage
- Key features of Live Lineage
- Introduction to Cloudera Lineage AI Assistant
- Key Capabilities of Cloudera Lineage AI Assistant
- Lab: Optimize & Validate Using Live Lineage & Cloudera Lineage AI Assistant

Managing Data Assets with the Knowledge Hub

- Introduction to the Knowledge Hub and Insight Dashboard
- Key Capabilities for Data Asset Management
- Collaboration and Governance Features
- Lab: Populating and Managing Assets in the Knowledge Hub