Training: Google Cloud
Managing a Data Mesh with Knowledge Catalog

Google Cloud

TRAINING GOALS:

Knowledge Catalog is an intelligent data fabric that enables organizations to centrally discover, manage, monitor, and govern their data across data lakes, data warehouses, and data marts. You can use Knowledge Catalog to build a data mesh architecture to decentralize data ownership among domain data owners. In this course, you will learn how to discover, manage, monitor, and govern your data across data lakes, data warehouses, and data marts through guided lectures and independent exercises using sample data.

What you'll learn:

- Secure data lakes, zones, and assets
- Implement tagging for resources and use tags to search for assets
- Process data using Knowledge Catalog tasks
- Design, execute and report on data quality processes

Audience

Every Knowledge Catalog user.

Products

- Knowledge Catalog
- BigQuery Studio
- Dataflow Prime
- Storage Insights
- Lakehouse
- Managed Service for Apache Spark

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

CONSPECT:

- Module 1 - Introduction to Knowledge Catalog
  - Topics:
    - Modern Data Platforms and Data-Oriented Design
    - Pillars of Data Governance
    - What is Knowledge Catalog?
    - Knowledge Catalog Capabilities
    - Knowledge Catalog compared with other products on Google Cloud
  - Objectives:
    - Identify the importance of a modern data platform
    - Explain the role of Knowledge Catalog on Google Cloud
- Module 2 - Creating a Data Mesh on Knowledge Catalog
  - Topics:
    - What is a data mesh?
    - Knowledge Catalog concepts
    - Creating data lakes and zones
    - Assets in Knowledge Catalog
  - Objectives:
    - Define key Knowledge Catalog concepts
    - Configure and set up Knowledge Catalog
  - Activities:
    - Lab: Provision a Data Mesh using Knowledge Catalog
- Module 3 - Processing Data on Knowledge Catalog
  - Topics:
    - Processing data on Knowledge Catalog
    - Data preparation tasks
    - Ingestion jobs
    - Dataflow and Spark tasks
  - Objectives:
    - Understand different data processing options in Knowledge Catalog
    - Configure and run data preparation tasks on Knowledge Catalog
- Module 4 - Managing Data Security through Knowledge Catalog
  - Topics:
    - IAM permissions and roles
    - Securing your data lake

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

        - Policy management
        - Metadata security
    - Objectives:
        - Secure data lakes, zones, and assets in Knowledge Catalog
- Module 5 - Data Tagging and Data Catalog
    - Topics:
        - Introduction to Data Catalog
        - Technical metadata vs. business metadata
        - Tags and tag templates
        - Entries and entry groups
        - Data lineage
    - Objectives:
        - Implement tagging for resources and use tags to search for assets
    - Activities:
        - Lab: Data Catalog and Data Lineage
- Module 6 - Data Quality and Profiling
    - Topics:
        - Data quality tasks and AutoDQ
        - Reporting on data quality
        - Data profiling
    - Objectives:
        - Design, execute and report on data quality processes
    - Activities:
        - Lab: Data Quality and Profiling your Data in BigQuery
- Module 7 - Knowledge Catalog Best Practices
    - Topics:
        - Best practices
        - End-to-end demo
    - Objectives:
        - Implement best practices for Knowledge Catalog
    - Activities:
        - Challenge Lab: Managing a Data Mesh with Knowledge Catalog

REQUIREMENTS:

Completion of the "Modernizing Data Lakes and Data Warehouses with Google Cloud" and "Building

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007

Batch Data Pipelines on Google Cloud" courses in the "Data Engineer" learning path or equivalent experience using Google Cloud.

Difficulty level



CERTIFICATE:

The participants will obtain certificates signed by Google Cloud (course completion).

TRAINER:

Authorized Google Cloud Trainer

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
EUR: PL 79 1600 1462 1853 6451 3000 0006
USD: PL 52 1600 1462 1853 6451 3000 0007