

## ***AW470 Building Data Lakes on AWS***

### **Kurzbeschreibung:**

In diesem Kurs lernen Sie das Erstellen eines Operationalen Data Lake für die Analyse von strukturierten und unstrukturierten Daten. Sie lernen die relevanten Komponenten und Funktionalitäten der für die Erstellung eines Data Lakes beteiligten Services kennen. Sie lernen den Umgang mit AWS Lake Formation um Data Lakes zu erstellen, AWS Glue zum Anlegen eines Data Catalog und Amazon Athena um Daten zu analysieren. Weiterhin lernen Sie im Kurs und im Umgang mit den Labs das Untersuchen von gewöhnlichen Data Lake Architekturen.

### **Zielgruppe:**

Der Kurs richtet sich an:

- Data platform engineers
- Solutions architects
- IT professionals

### **Voraussetzungen:**

Um an dem Kurs „Building Data Lakes on AWS“ bei qSkills teilnehmen zu können, sollten Sie die folgenden AWS-Trainings besucht haben:

- „Data Analytics Fundamentals“  
oder ein Jahr Erfahrung im Erstellen von Data Analytics Pipelines
- „AWS Technical Essentials“

### **Sonstiges:**

**Dauer:** 1 Tage

**Preis:** 750 Euro plus Mwst.

### **Ziele:**

In diesem Kurs lernen Sie:

- Die Anwendung relevanter Methoden für die Planung und das Design von Data Lakes
- Sie kennen die Komponenten und Services, die für das Anlegen von Data Lakes relevant sind
- Das Sichern von Data Lakes mit entsprechenden Zugriffsrechten
- Aufnehmen, Speichern und Transformieren von Daten in einem Data Lake
- Abfragen, Analysieren und Visualisieren von Daten innerhalb eines Data Lake

## Inhalte/Agenda:

- - ◆ Modul 1: Einführung in Data Lakes
    - ◆ ◇ Der Wert von Data Lakes
    - ◆ ◇ Data Lakes vs. Data Warehouses
    - ◆ ◇ Komponenten von Data Lakes
    - ◆ ◇ Allgemeine Architekturen aufbauend auf Data Lakes
  - ◆ Modul 2: Datenaufnahme, Katalogisierung und Vorbereitung
    - ◆ ◇ Beschreiben der Beziehung zwischen Datenaufnahme und Datenspeicherung
    - ◆ ◇ AWS Glue crawlers und wie sie genutzt werden, um einen Datenkatalog zu erstellen
    - ◆ ◇ Identifizieren der Datenformatierung, Partitionierung und Kompression für eine effiziente Speicherung und Abfrage
    - ◆ ◇ Lab 1: Aufsetzen eines einfachen Data Lake
  - ◆ Modul 3: Datenverarbeitung und -Analyse
    - ◆ ◇ Datenverarbeitung bei Data Lakes
    - ◆ ◇ Anwendung von AWS Glue für die Verarbeitung von Daten innerhalb von Data Lakes
    - ◆ ◇ Nutzung von Amazon Athena für die Analyse von Daten in einem Data Lake
  - ◆ Modul 4: Erstellen eines Data Lake mit AWS Lake Formation
    - ◆ ◇ Nutzen und Features von AWS Lake Formation
    - ◆ ◇ Anwendung von AWS Lake Formation zur Erstellung eines Data Lake
    - ◆ ◇ Das AWS Lake Formation Sicherheitsmodell
    - ◆ ◇ Lab 2: Erstellen eines Data Lake mit AWS Lake Formation
  - ◆ Modul 5: Zusätzliche Lake Formation Konfigurationen
    - ◆ ◇ Automatisierung von AWS Lake Formation mit blueprints und workflows
    - ◆ ◇ Anwendung von Sicherheits- und Zugangskontrolle zu AWS Lake Formation
    - ◆ ◇ Abstimmung von Datensätzen mit AWS Lake Formation FindMatches
    - ◆ ◇ Visualisierung von Daten mit Amazon QuickSight
    - ◆ ◇ Lab 3: Data Lakes automatisiert erstellen mit AWS Lake Formation blueprints
    - ◆ ◇ Lab 4: Datenvisualisierung mit Amazon QuickSight
  - ◆ Modul 6: Architektur und Kursrückblick
    - ◆ ◇ Wissenscheck
    - ◆ ◇ Architektur
    - ◆ ◇ Kursrückblick