

## Big Data - Data Engineering Grundlagen (54010)

ID 54010 Preis auf Anfrage Dauer 3 Tage

### Zielgruppe

Das Seminar richtet sich an Datenbank Experten, BigData Experten, Sicherheitsexperten, Systemadministratoren, Netzwerkadministratoren, Data Science Mitarbeiter, Entwickler sowie alle Interessierten, die dem Thema BigData tiefer unter die Haube schauen möchten.

### Voraussetzungen

Neben generellen IT Kenntnissen in SQL und Programmierung sind keine besonderen Vorkenntnisse erforderlich. Der vorherige Besuch unseres Grundlagenseminars Big Data – Definitionen, Technologien und Produkte im Überblick erleichtert das Verständnis.

### Kursinhalt

#### Data Engineering

#### Einführung Big Data:

- Was ist Big Data?
- Die drei Vs
- Abgrenzung zu relationalen Datenbanken
- Data Lake und hybride Architektur Konzept

#### IT Grundlagen:

- Basiswissen Netzwerke
- Basiswissen verteilte Systeme

#### Ökosystem Hadoop:

- Hadoop Cluster Architektur
- Überblick über die Hadoop Distributionen
- HBase
- HDFS
- Hive
- Kafka
- NiFi
- Pig
- Spark / Spark Streaming
- Sqoop

- Ambari
- Zookeeper
- Überblick weiterer Komponenten

#### Security in Hadoop:

- Ranger
- Kerberos

#### Hands on Hadoop – Praxisteil:

- HDFS Puzzle
- Administration Grundlagen
- Ambari Monitoring M/R
- HDFS Strukturen anlegen
- Zookeeper Nodes
- HBase
- HDFS
- Hive
- Kafka
- NiFi
- Pig
- Spark / Spark Streaming
- Sqoop

#### Vertiefung der Hadoop Komponenten - Praxisteil:

- Streaming
- Buffering
- Archivierung
- Transformation
- Analyse

## Weltweite Trainingscenter



## Fast Lane Institute for Knowledge Transfer (Switzerland) AG

Husacherstrasse 3  
CH-8304 Wallisellen  
Tel. +41 44 832 50 80

info@flane.ch, <https://www.flane.ch>