


Data Engineer Training

Dreitägiges Intensivseminar



Mehr als 25 Jahre Kompetenz und Erfahrung –
lernen Sie uns kennen.
Herzlich willkommen bei saracus!

Dieses dreitägige Seminars gibt Data-Engineers einen tiefen Einblick in den Bereich Big Data Applikations-Entwicklung. Hierbei werden bekannte und weit verbreitete Frameworks wie Spark und Hadoop eingesetzt und zusätzlich verschiedene Tools wie Hive, Sqoop und viele weitere aus dem Hadoop Ökosystem verwendet.

Das Seminar soll verschiedene Methoden vermitteln, wie die Vielzahl an Tools sinnvoll miteinander verknüpft werden können und die Teilnehmer vorbereiten, die richtige Toolauswahl für vorhandene Problemstellungen treffen zu können.

Der Fokus dieses Seminars liegt darin, den Data-Engineers das nötige Wissen zu vermitteln, wie Daten in ein Hadoop-Cluster geladen und dort gespeichert werden, wie diese Daten in einem verteilten System skalierbar transformiert werden und wie mit Spark und SparkML eine Pipeline für Machine-Learning-Algorithmen aufgebaut werden kann.

Alle Übungen werden auf einem Multi-Node Hadoop Cluster ausgeführt, welches in der AWS-Cloud installiert ist. Als Hadoop-Distribution wird Cloudera verwendet.

An wen richtet sich das Seminar?

Das Seminar richtet sich an Software-Engineers, die verteilte Applikationen verstehen und entwickeln möchten, die auf einem Hadoop Cluster laufen. Dieses Seminar baut auf dem Einführungsseminar für Big Data auf.

Voraussetzungen

Notwendig ist ein Basiswissen von Python, Java und der Linux Shell.

Agenda - Tag 1 (Introduction to Core Hadoop)

09.00 Uhr

Begrüßung durch die Seminarleitung

09.15 Uhr – 17.00 Uhr

- Advanced HDFS
- Übung: Interaktion mit dem HDFS über CLI und GUI
- YARN internals
- Übung: Job debugging im Resource Manager
- Nutzung des Java MapReduce Frameworks für parallele Datenverarbeitung
- Übung: ETL Jobs mit dem MapReduce Framework
- Hadoop File Formats

Agenda - Tag 2 (Batch Analytics)

- Data Ingestion mit Apache Sqoop
- Übung: Datenextraktion aus einer relationalen Datenbank
- Apache Hive
- Übung: Hive Basics
- Advanced Hive
- Übung: Hive Advanced

Agenda - Tag 3 (Spark)

- Spark Core Framework
- Übung: ETL mit Apache Spark
- Spark SQL und Dataframes
- Übung: Interactive analytics mit Spark SQL
- SparkMLlib
- Übung: Machine Learning mit SparkML

Kaffee- und Teepausen finden nach Bedarf statt!

Anmeldung zum Seminar



Ja, ich nehme an dem Seminar
„Data Engineer Training (dreitägig)“ teil.

Firma: _____

Name: _____ Vorname: _____

Funktion: _____

Straße: _____

PLZ/Ort: _____

Tel.-Nr.: _____ Fax-Nr.: _____

Email: _____

Datum: _____ Unterschrift: _____

**Die Seminargebühr für dieses dreitägige Seminar beträgt 2.610 Euro in Deutschland /
2.990 CHF in der Schweiz pro Teilnehmer.**

Gebuchter Seminartermin (bitte angeben): _____

Ihre Anmeldung sollte spätestens 14 Tage vor dem Seminartermin erfolgen.

Gebuchter Seminarort (bitte kreuzen Sie an):

Baden-Dättwil(CH)

Münster (D)

Ich benötige ein Hotel!

Nach Ablauf der Anmeldefrist erhalten Sie von uns eine Anmeldebestätigung und eine Rechnung sowie eine Anfahrtsskizze. Bitte überweisen Sie den Rechnungsbetrag vor Veranstaltungsbeginn.

**Bitte senden Sie diese Seite per Fax an +49 (0251) 98721-26
oder per Post an: saracus consulting GmbH
„Data Engineer Training (dreitägig)“
Hafenweg 46, D- 48155 Münster**

Voraussetzung:

Berücksichtigen Sie bitte, dass eine Mindestteilnehmerzahl von 4 Personen vorgesehen ist. Ansonsten kann das Seminar nicht stattfinden.

Anmeldung:

Bitte nehmen Sie Ihre Abmeldung schriftlich vor. Selbstverständlich ist eine Vertretung des angemeldeten Teilnehmers möglich. Alle genannten Preise verstehen sich zzgl. der gesetzl. MwSt.

Stornierung:

Bei Stornierung bis zu 14 Tage vor Veranstaltungsbeginn erheben wir eine Bearbeitungsgebühr von 50% der Gebühr. Bei späteren Absagen wird der gesamte Beitrag fällig.

Seminarabsage:

Sollten wir aus wichtigen Grund (z. B. Erkrankung des/der Referenten) gezwungen sein, den Kurs abzusagen, so teilen wir Ihnen dieses umgehend mit. Wir werden Ihnen in diesem Fall einen Ersatztermin anbieten. Passt dieser nicht zu Ihrer Terminplanung, erhalten Sie die bereits gezahlte Seminargebühr in voller Höhe zurück. Darüber hinausgehende Ansprüche bestehen nicht.



**Für Rückfragen steht Ihnen gerne das Team der Data & Analytics academy der saracus consulting unter +49 (0251) 98721-0 zur Verfügung.
Email: saracus@saracus.com
Internet: <https://www.saracus.com/know-how/data-and-analytics-academy/>**