

Big Data in Unternehmen

Alles Wichtige auf einen Blick

Entdecken Sie die Macht Ihrer Daten! In der digitalen Wirtschaft von heute entscheidet die Fähigkeit zur Datennutzung über Erfolg oder Misserfolg. Dieser Überblicks-Kurs zu Big Data öffnet Ihnen die Türen zur faszinierenden Welt von Big Data – auch ohne Programmiererfahrung!

Sie lernen die Grundlagen der wichtigsten Technologien wie Hadoop, Spark, Elasticsearch und Kafka kennen und verstehen, wie moderne Unternehmen aus riesigen Datenmengen wertvolle Geschäftserkenntnisse gewinnen. Durch anschauliche Beispiele und interaktive Demonstrationen entwickeln Sie ein solides Verständnis für Big Data-Infrastrukturen, Datenpipelines und Real-Time-Analytics. Sie erfahren, wie der ELK-Stack für Monitoring und Visualisierung eingesetzt wird und wie Cloud-Plattformen skalierbare Datenlösungen ermöglichen.

Mit Low-Code-Tools und visuellen Workflows lernen Sie, selbst komplexe Datenanalysen durchzuführen und aussagekräftige Dashboards zu erstellen. Dieses Big Data Training verwandelt Sie vom Daten-Betrachter zum Data-Driven Decision Maker. Sie gewinnen das Vertrauen, Big Data-Projekte zu verstehen, zu bewerten und erfolgreich zu begleiten – ein unverzichtbarer Wettbewerbsvorteil in der modernen Geschäftswelt. Erwerben Sie die Fähigkeiten, um die datengetriebene Zukunft Ihres Unternehmens mitzugestalten!

Kursinhalt

- Was ist Big Data?
- Die 4 V's und Abgrenzung zu klassischer Datenverarbeitung
- Anwendungsbeispiele, Umsetzung und Praktische Einsatzgebiete
- Low-Code-/No-Code-Analytics mit KNIME
- Big Data Visualisierung und Operational Intelligence mit Kibana und PowerBI
- MapReduce und CAP-Theorem anhand der Hadoop Architektur und HDFS-Konzepte
- Apache Spark und In-Memory-Verarbeitung
- Datenstreaming und Echtzeitverarbeitung
- Elasticsearch und ELK-Stack
- Big Data in der Cloud mit AWS, Azure, Google Cloud und Databricks
- Data-Lakes vs. Data-Warehouses
- Effiziente Datenpipelines mit Logstash, Beats und Apache Airflow
- Data Governance und Compliance
- Zukunftstrends und neue Technologien

E-Book Das ausführliche deutschsprachige digitale Unterlagenpaket, bestehend aus PDF und E-Book, ist im Kurspreis enthalten.

Zielgruppe

Dieser Kurs richtet sich an alle, die den Einstieg in die Welt der Big-Data-Lösungen suchen – sei es zur Planung, Bewertung oder praktischen Umsetzung. Ob Sie erste Ideen entwickeln oder bestehende Strategien optimieren möchten: Hier erhalten Sie das nötige Wissen und die Werkzeuge, um datengetriebene Projekte erfolgreich voranzubringen.

Voraussetzungen

Neugier und Interesse an der Arbeit mit großen Datenmengen sind wichtig! Darüber hinaus ist ein grundlegendes Verständnis von Datenstrukturen und -verarbeitung hilfreich.

Kursziel

Sie verstehen Big-Data-Grundlagen, erkennen relevante Anwendungsfälle und beherrschen einen praxisnahen Workflow von Datenerfassung über Verarbeitung bis Visualisierung und Deployment – inklusive Datenqualität und Skalierung.

Dieser Kurs im Web



Alle tagesaktuellen Informationen und Möglichkeiten zur Bestellung finden Sie unter dem folgenden Link: www.experteach.de/go/BIGD

Vormerkung

Sie können auf unserer Website einen Platz kostenlos und unverbindlich für 7 Tage reservieren. Dies geht auch telefonisch unter 06074 4868-0.

Garantierte Kurstermine

Für Ihre Planungssicherheit bieten wir stets eine große Auswahl garantierter Kurstermine an.

Ihr Kurs maßgeschneidert

Diesen Kurs können wir für Ihr Projekt exakt an Ihre Anforderungen anpassen.

Training		Preise zzgl. MwSt.	
Termine in Deutschland	3 Tage	€ 2.195,-	
Online Training	3 Tage	€ 2.195,-	
Termin/Kursort	Kursprache Deutsch 		
27.05.-29.05.26  München	26.10.-28.10.26  Düsseldorf		
27.05.-29.05.26  Online	26.10.-28.10.26  Online		

Stand 05.03.2026

Inhaltsverzeichnis

Big Data in Unternehmen – Alles Wichtige auf einen Blick

1 Was ist Big Data?	4.9 Flink	8.5 Datenrisiken
1.1 Der große Berg Daten	4.10 Oozie	8.5.1 Sicherheit der Daten
1.2 Anwendungsgebiete von BIG DATA	4.11 Zookeeper	8.5.2 Qualität der Daten
1.3 Die Definition von Big Data: 3-5 „V“s	4.12 Ambari	8.6 Definitions- und Aussagerisiko
1.3.1 Volume	4.13 R Connector	8.6.1 Faktor Daten
1.3.2 Velocity: Geschwindigkeit und Aktualität der Daten	4.14 Cassandra	8.6.2 Faktor Data Management
1.3.3 Variety: Vielfältigkeit der Daten	4.15 SAP HANA	8.6.3 Faktor Organisation
1.3.4 Veracity: Gültigkeit und Zuverlässigkeit der Daten		8.6.4 Faktor Prozess
1.3.5 Das ist Big Data		8.6.5 Faktor Kunde als Betroffener
1.4 Der Ursprung von Big Data		8.7 Herausforderungen
		8.8 EU-AI-Act: Erste Regulierungen von KI
2 Big Data Grundlagen	5 Anforderungen an die Datenbankentwicklung	9 Herausforderungen im Betrieb von Big Data Lösungen
2.1 Die BIG DATA Value Chain	5.1 Die Entwicklung von SQL zu NoSQL	9.1 Wo ist der Anfang?
2.2 Quellen für BigData Datenanalysen	5.2 Facebook und Twitter: höhere Anforderungen	9.2 Hadoop unternehmensweit betreiben
2.3 Die Architektur	5.2.1 Read-Only-Replicas zur Überbrückung	9.2.1 Physische Infrastruktur
2.4 SQL: feste, vordefinierte Tabellenschemata	5.2.2 Multi-Master Datenbanken	9.2.2 Datenhaltung
2.5 Normalisierung von Tabellen	5.3 ACID vs BASE	9.2.3 Datenzugriff
2.6 NoSQL	5.4 Vorteile von NoSQL (BASE)	9.2.4 Datenintegration
2.6.1 Key-Value Stores	5.5 Der ELK Stack	9.2.5 IT-Sicherheit
2.6.2 In-Memory Key-Value Stores	5.5.1 Elasticsearch Abfragen auf der Konsole	9.2.6 Weitere Betriebskriterien
2.6.3 Document Stores	5.5.2 Kibana - das Dashboard für grafische Anzeigen	9.2.7 Wirtschaftliche Kriterien
2.6.4 Graph Databases	5.6 ELK: Sharding und Reverse Indexing	9.3 Betriebsphasen einer neuen Big-Data Landschaft
2.6.5 Column Stores	5.7 MongoDB - ein beliebter Document Store	9.4 Erfolgsfaktoren
2.7 CAP-Theorem		
2.7.1 Kombination CA des CAP-Theorems	6 Anwendungsbeispiele für Big Data	10 Ausblick
2.7.2 Kombination CP des CAP-Theorems	6.1 Limitierungen klassischer analytischer Anwendungen	10.1 Prognose der Big Data Markt-Kapitalisierung
2.7.3 Kombination AP des CAP-Theorems	6.2 Einsatzszenarien von Big Data	10.2 Marktanteile führender Unternehmen im 4. Quartal 2023
	6.2.1 Clickstream Analyse	10.3 Orchestrierung in der Cloud
	6.2.2 Stimmungsanalysen aus Social Media	10.4 KI und ML
	6.2.3 Analyse von Log-Daten	10.5 Environment-, Social- und Governance-Faktoren (ESG)
	6.2.4 Analyse von Sensordaten	10.6 Quantencomputing und Big Data
	6.2.5 Analyse von Texten	10.7 Natural Language Processing (NLP)
	6.2.6 Analyse von Video- und Sprachdaten	10.8 Big-Data- und ML/KI Fachkräftebedarf
	6.3 Weltweite Immobilienvermietung am Beispiel von Airbnb	10.9 Aktueller Stand
	6.4 Internationale Datacenterverknüpfung am Beispiel CERN	10.10 Technische Entwicklungen
	6.5 Beispiel: US FINRA	10.11 Markt-Entwicklungen
	6.6 Beispiel: Jobbörse MONSTER	10.12 Business-Entwicklungen
	6.7 Beispiel: Google BigTable	10.13 Diskussion der Ergebnisse
	7 Pipelines und ETL Prozesse	11 Hadoop Installation & Configuration & Go!
	7.1 DB vs. Datawarehouse vs. Data Lake	11.1 Installationsschema für APACHE HADOOP 3.1.3
	7.2 Snowflake-Schemas	11.2 Hadoop 3.1.3 auf Github
	7.3 Extract - Transform - Load (ETL)	11.3 Die Expertech Labor-Umgebung
	7.4 Pivot Tabellen / Klassische Reports	11.4 Anpassung der Konfigurationsdateien
	7.5 Pipeline für einen ETL Prozess	11.5 Übersicht der Dateistrukturen im Labor
	7.6 Data-Lakes	11.6 Erster Start des HDFS
		11.7 Syntax und Ablauf von Zählaufgaben
		11.8 Output beim MAPREDUCE Vorgang
		11.9 Hadoop Cockpit
		11.10 Wordcount Abfrage via PIG:
		11.11 RATING – Datensätze filtern (25 Mio)
4 Big Data Technologien	8 Data Governance + Risiken	
4.1 Das Hadoop Ecosystem	8.1 Die 3 Säulen von Data Governance	
4.2 Pig	8.2 Was kann ich tun, um meine Daten zu schützen?	
4.3 Hive	8.3 Risiken	
4.4 Mahout	8.4 Data Compliance Risiko	
4.5 HBase	8.4.1 National und in Europa	
4.6 Sqoop	8.4.2 Inhalte der DSGVO	
4.7 Flume	8.4.3 International	
4.8 Chukwa	8.4.4 Sozialrisiko	

