

Modulbeschreibung „Data Architecture – Wissen und Praxis für Data Engineering & KI“

Titel des Zertifikatskurses	Data Architecture – Wissen und Praxis für Data Engineering & KI
Kürzel/Modulnummer	---
Fachbereich	03 Elektrotechnik und Informatik
Modulverantwortliche/r	Prof. Dr. Christoph Quix, christoph.quix@hs-niederrhein.de
Dozent/in	Prof. Dr. Christoph Quix Prof. Dr. Timo Schwarzwälder
Modultyp	Hochschulzertifikatskurs der WWB
Dauer	Ca. 5 Termine in 2 Monaten
Häufigkeit des Angebots	Voraussichtlich jährlich und auf Nachfrage (Inhouse)
Zielgruppe(n)	Der Zertifikatskurs richtet sich an Fach- und Führungskräfte aller Branchen aus den Bereichen Informationsmanagement, Organisation und Prozessmanagement... <ul style="list-style-type: none"> • ...die die strategische Planung von Informationsarchitekturen in einem Unternehmen verantworten. • ...die Informationssysteme und -prozesse in Unternehmen definieren und deren Einsatz koordinieren. • ... die Architekturen und Systeme für Big Data- und Data Science-Anwendungen planen und konzipieren.
Angestrebte Lernergebnisse/ Learning outcomes	Mit erfolgreichem Abschluss des Kurses werden die Teilnehmenden in der Lage sein, <ul style="list-style-type: none"> • die wesentlichen Data Governance-Konzepte in konkreten Anwendungsfällen anzuwenden. • Data Privacy und Data Compliance in Big Data-Projekten einzuhalten und die notwendigen rechtlichen Vorgaben (z.B. DSGVO) zu kennen. • verschiedene Architekturen von Data Management-Systemen zu erklären und zu vergleichen. • die Datenmodelle von diversen NoSQL- bzw. Big Data- Systemen zu erklären und für die Datenmodellierung anzuwenden. • die Methoden des Data Engineering zur Aufbereitung, Transformation und Integration von Daten zu erläutern und entsprechende Systeme zur Umsetzung zu beurteilen. • die Qualität von Daten für datenbasierte Anwendungen zu beurteilen und grundlegende Konzepte des Datenqualitätsmanagements einzusetzen. • die Vor- und Nachteile von verteilten Big Data-Systemen zu beurteilen und deren Einsatz zu begründen. • Datentransformationen und Datenabfragen in Big Data-Systemen zu konzipieren. • Big Data-Systeme für bestimmte Anforderungen zu entwerfen. • eigene oder existierende Big Data-Systeme zu evaluieren.

Inhalte	<ul style="list-style-type: none"> • Grundlagen von relationalen und NoSQL-Datenbank-Systemen • Data Governance und ihre Konzepte • Datenqualitätsmanagement, sowohl organisatorisch als auch die technische Umsetzung zur Messung und Sicherung von Datenqualität • Data Management-Architekturen (Data Warehouse, Data Lake, Data Mesh, Data Fabric, usw.) • Datenaufbereitung, Datenintegration und Datenbereinigung • Abfragesprachen für verschiedene Datenbanksysteme • Grundlagen und Architekturen Big Data-Systeme (z.B. Hadoop, Spark, Kafka) • Konzepte des verteilten Data Managements • Herausforderungen bei Big Data-Integration
Lehrformen	<p>Blended-Learning-Ansatz als Kombination von Präsenzveranstaltungen und E-Learning-Inhalten</p> <ul style="list-style-type: none"> • Grundlegende Informationen zum Angleichen des Wissensstands der Teilnehmenden im Vorfeld der ersten Präsenzveranstaltung als E-Learning-Kurs • Präsenzveranstaltung mit der Präsentation von theoretischen Grundlagen und praktischen Übungen an verschiedenen Data Management-Systemen auf der Basis von Fallstudien • Projektarbeit als Hausaufgabe zur Vertiefung der erlernten Inhalte • Abschließende Präsenzveranstaltung mit Diskussion der Projektarbeiten und weiterführender Inhalte; Reflektion und Zusammenfassung der gelernten Inhalte
Unterrichtssprache	Deutsch
Teilnahmevoraussetzungen	<p>Hochschulabschluss oder anderweitiger berufsqualifizierender Abschluss mit einschlägiger Berufserfahrung.</p> <p>Grundlegende Kenntnisse von Datenarchitekturen und der Datenmodellierung sollten vorhanden sein.</p>
Abschluss	Hochschulzertifikat (Prüfungsteilnahme) oder Teilnahmebescheinigung (75% Anwesenheit)
Prüfungsleistung(en)	Projektarbeit mit schriftlicher Dokumentation und Vortrag
Leistungspunkte	3 ECTS
Workload/Arbeitsaufwand	75 h Gesamtstunden
Kontaktzeit	40 h Präsenz
Selbststudium	35 h (Nachbereitung und Prüfungsvorbereitung, Erstellung einer Projektarbeit etc.)
Geplante Gruppengröße	ca. 12 Teilnehmende
Verwendbarkeit des Moduls	---
Literatur	<p>Hinweise zur Basisliteratur</p> <ul style="list-style-type: none"> • M. Kleppmann: Designing Data-Intensive Applications. O'Reilly, 2017. Auch in Deutsch verfügbar: Datenintensive Anwendungen designen. O'Reilly, 2019.

	<ul style="list-style-type: none">• J. Reis, M. Housley: Fundamentals of Data Engineering. O'Reilly, 2022• P.J. Sadalage, M. Fowler: NoSQL Distilled. Addison-Wesley, 2012.• G. Harrison. Next Generation Databases. Apress, 2015.
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------