

## Modulbeschreibung „Data Architecture & Data Engineering“

Titel des Zertifikatskurses	Data Architecture & Data Engineering
Kürzel/Modulnummer	---
Fachbereich	03 Elektrotechnik und Informatik
Modulverantwortliche/r	Prof. Dr. Christoph Quix, <a href="mailto:christoph.quix@hs-niederrhein.de">christoph.quix@hs-niederrhein.de</a>
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"> <li>• ...die die strategische Planung von Informationsarchitekturen in einem Unternehmen verantworten.</li> <li>• ...die Informationssysteme und -prozesse in Unternehmen definieren und deren Einsatz koordinieren.</li> <li>• ... die Architekturen und Systeme für Big Data- und Data Science-Anwendungen planen und konzipieren.</li> </ul>
Angestrebte Lernergebnisse/ Learning outcomes	Mit erfolgreichem Abschluss des Kurses werden die Teilnehmenden in der Lage sein, <ul style="list-style-type: none"> <li>• die wesentlichen Data Governance-Konzepte in konkreten Anwendungsfällen anzuwenden.</li> <li>• Data Privacy und Data Compliance in Big Data-Projekten einzuhalten und die notwendigen rechtlichen Vorgaben (z.B. DSGVO) zu kennen.</li> <li>• verschiedene Architekturen von Data Management-Systemen zu erklären und zu vergleichen.</li> <li>• die Datenmodelle von diversen NoSQL- bzw. Big Data- Systemen zu erklären und für die Datenmodellierung anzuwenden.</li> <li>• die Methoden des Data Engineering zur Aufbereitung, Transformation und Integration von Daten zu erläutern und entsprechende Systeme zur Umsetzung zu beurteilen.</li> <li>• die Qualität von Daten für datenbasierte Anwendungen zu beurteilen und grundlegende Konzepte des Datenqualitätsmanagements einzusetzen.</li> <li>• die Vor- und Nachteile von verteilten Big Data-Systemen zu beurteilen und deren Einsatz zu begründen.</li> <li>• Datentransformationen und Datenabfragen in Big Data-Systemen zu konzipieren.</li> <li>• Big Data-Systeme für bestimmte Anforderungen zu entwerfen.</li> <li>• eigene oder existierende Big Data-Systeme zu evaluieren.</li> </ul>

Inhalte	<ul style="list-style-type: none"> <li>• Grundlagen von relationalen und NoSQL-Datenbank-Systemen</li> <li>• Data Governance und ihre Konzepte</li> <li>• Datenqualitätsmanagement, sowohl organisatorisch als auch die technische Umsetzung zur Messung und Sicherung von Datenqualität</li> <li>• Data Management-Architekturen (Data Warehouse, Data Lake, Data Mesh, Data Fabric, usw.)</li> <li>• Datenaufbereitung, Datenintegration und Datenbereinigung</li> <li>• Abfragesprachen für verschiedene Datenbanksysteme</li> <li>• Grundlagen und Architekturen Big Data-Systeme (z.B. Hadoop, Spark, Kafka)</li> <li>• Konzepte des verteilten Data Managements</li> <li>• Herausforderungen bei Big Data-Integration</li> </ul>
Lehrformen	<p>Blended-Learning-Ansatz als Kombination von Präsenzveranstaltungen und E-Learning-Inhalten</p> <ul style="list-style-type: none"> <li>• Grundlegende Informationen zum Angleichen des Wissensstands der Teilnehmenden im Vorfeld der ersten Präsenzveranstaltung als E-Learning-Kurs</li> <li>• Präsenzveranstaltung mit der Präsentation von theoretischen Grundlagen und praktischen Übungen an verschiedenen Data Management-Systemen auf der Basis von Fallstudien</li> <li>• Projektarbeit als Hausaufgabe zur Vertiefung der erlernten Inhalte</li> <li>• Abschließende Präsenzveranstaltung mit Diskussion der Projektarbeiten und weiterführender Inhalte; Reflektion und Zusammenfassung der gelernten Inhalte</li> </ul>
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"> <li>• M. Kleppmann: Designing Data-Intensive Applications. O'Reilly, 2017.</li> </ul> <p>Auch in Deutsch verfügbar: Datenintensive Anwendungen designen. O'Reilly, 2019.</p>

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