Modulbeschreibung „Data Engineering – Datenarchitekturen gestalten“

Titel des Zertifikatskurses	Data Engineering – Datenarchitekturen gestalten
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. 3 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… • …die die strategische Planung von Informationsarchitekturen in einem Unternehmen verantworten. • …die Informationssysteme und -prozesse in Unternehmen definieren und deren Einsatz koordinieren.
Angestrebte Lernergebnisse/ Learning outcomes	Mit erfolgreichem Abschluss des Kurses werden die Teilnehmenden in der Lage sein, • Die Anwendungsbereiche und Einschränkungen von klassischen relationalen Datenbanksystemen zu verstehen. • 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. • Datentransformationsprozesse zur Aufbereitung und Bereinigung von Daten zu implementieren. • Datenabfragen in verschiedenen Datenbanksystemen zu definieren.
Inhalte	• Grundlagen: Daten, Informationen; Datenmodellierung; Datenbanksysteme • Vorgehensmodelle zur Datenanalyse (CRISP-DM, Knowledge Discovery Pipeline) und die Rolle des Data Engineering; Relationale und NoSQL-Datenbank-Systeme; Konzeptuelle und logische Datenmodellierung; Übungen (Reverse Engineering von Datenmodellen, Datenmodellierung mit UML, Vorgehen bei der Datenmodellierung); Umsetzung von Datenmodellen am Rechner • Vor- und Nachteile von NoSQL- und relationalen Systemen; Rolle der Datenmodellierung im Unternehmen

	• Datenmanagement-Architekturen: Data Warehouse und Data Lake; Lambda-Architektur und andere Ansätze zur effizienten Verarbeitung von Big Data • Interaktion mit Datenbank-Systemen: Definition der notwendigen Datenbank-Strukturen; Strukturierung und Import von Daten; Abfragesprachen, insbesondere für SQL- und NoSQL-Systeme (z. B. MongoDB und Neo4j) • Vergleich und Bewertung der Datenbank-Systeme**:** Kriterien zur Beurteilung von Datenbank-Systemen, Vergleich und Bewertung anhand von Fallstudien
Lehrformen	Blended-Learning-Ansatz als Kombination von Präsenzveranstaltungen und E-Learning-Inhalten • Grundlegende Informationen zum Angleichen des Wissenstands 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	Hochschulabschluss oder anderweitiger berufsqualifizierender Abschluss mit einschlägiger Berufserfahrung. Grundlegende Kenntnisse von Datenarchitekturen und der Datenmodellierung sollten vorhanden sein.
Abschluss	Hochschulzertifikat (Prüfungsteilnahme) oder Teilnahmebescheinigung (75% Anwesenheit)
Prüfungsleistung(en)	Projektarbeit mit schriftlicher Dokumentation und Vortrag
Leistungspunkte	3 ECTS
Workload/Arbeitsaufwand	75 h Gesamtstunden
Kontaktzeit	24 h Präsenz
Selbststudium	51 h (Nachbereitung und Prüfungsvorbereitung, Erstellung einer Projektarbeit etc.)
Geplante Gruppengröße	ca. 12 Teilnehmende
Verwendbarkeit des Moduls	---
Literatur	Hinweise zur Basisliteratur • M. Kleppmann: Designing Data-Intensive Applications. O‘Reilly, 2017. Auch in Deutsch verfügbar: Datenintensive Anwendungen designen. O‘Reilly, 2019. • G. Harrison. Next Generation Databases. Apress, 2015.

	• 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.