



Hochschule Niederrhein  
University of Applied Sciences

WISSENSCHAFTLICHE WEITERBILDUNG AN DER HOCHSCHULE NIEDERRHEIN

## Big Data Technologien

### Skalierbare, verteilte Big Data Systeme entwickeln

#### KURSBESCHREIBUNG

Zur Bewältigung der steigenden Informationsflut wurden in den letzten Jahren unter dem Schlagwort „Big Data“ viele Technologien auf den Markt gebracht, um effiziente, skalierbare Anwendungen zu entwickeln, die große Datenmengen verarbeiten können. Jedoch ist das Angebot an kommerzieller oder auch kostenloser Open-Source-Software so groß, dass es selbst Experten schwerfällt, geeignete Technologien für einen Anwendungsfall auszuwählen.

In diesem Zertifikatskurs erhalten Sie einen Überblick über verschiedene Big Data Systeme (z. B. Hadoop, Spark) und können mit praktischen Fallstudien lernen, für welche Anwendungsfälle diese geeignet sind. Darüber hinaus werden Ihnen die grundlegenden Prinzipien von verteilten, skalierbaren Big Data Architekturen vermittelt, um Systeme entsprechend einordnen und bewerten zu können. Durch praktische Übungen mit aktuellen Big Data Systemen und der Diskussion von individuellen Fragestellungen können Sie die Inhalte des Kurses in Ihren Arbeitsalltag übertragen.

#### KURSZIELE

Mit erfolgreichem Abschluss des Kurses werden Sie in der Lage sein:

- Verschiedene Architekturen von Big Data Systemen zu erklären und zu vergleichen.
- Die Vor- und Nachteile von verteilten Big Data Systemen zu beurteilen und deren Einsatz zu begründen.
- Datentransformationen und Datenabfragen in Big Data Systemen zu implementieren.
- Einfache Big Data Systeme für bestimmte Anforderungen zu entwerfen und zu entwickeln.
- Eigene oder existierende Big Data Systeme zu evaluieren.

#### VORTEILE

- Wissens- und Kompetenzzuwachs durch die Vermittlung von zeitgemäßen Big Data Architekturen und -Systemen aus Forschung und Praxis.
- Unabhängige und kritische Diskussion und Bewertung aktueller Big Data Technologien.
- Berufliche Verwertbarkeit der vermittelten Inhalte durch praxisnahe Übungen und Projektarbeiten.
- Dieser Zertifikatskurs ist Teil des Zertifikatsstudiums:

#### CAS Data Architect

Informationsarchitekturen gestalten - Daten effizient verwalten

Das Zertifikatsstudium umfasst insgesamt drei Zertifikatskurse. Nach erfolgreichem Absolvieren erhalten Sie das Certificate of Advanced Studies (CAS) der Hochschule Niederrhein:

„Data Architect“

Mehr Informationen zum CAS finden Sie hier:  
<https://www.hs-niederrhein.de/weiterbildung/data-architect/>

KOSTENÜBERNAHME DURCH DAS VOM BMBF GEFÖRDERTE PROJEKT

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung



## PRÄSENZTERMINE

Präsenz 1: MO, 30.03.2020 | 9–17 Uhr

Präsenz 2: DI, 31.03.2020 | 9–17 Uhr

Präsenz 3: FR, 08.05.2020 | 9–17 Uhr

Vor und zwischen den Präsenztagen liegen onlinegestützte Selbstlernphasen.

## 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.

## TEILNAHMEVORAUSSETZUNGEN

Hochschulabschluss mit mindestens einjähriger Berufserfahrung oder anderweitiger berufsqualifizierender Abschluss mit mindestens dreijähriger Berufstätigkeit.

Grundlegende Kenntnisse von Datenarchitekturen sollten vorhanden sein. Zwingend erforderlich ist ein eigenes Notebook mit der Möglichkeit, Programme zu installieren und auszuführen.

## LEHR- UND LERNFORM

Der in einem interaktiven Seminarcharakter gehaltene Kurs bietet die Möglichkeit, auf individuelle Frage- und Problemstellungen der Teilnehmenden einzugehen. Praktische Umsetzungsaufgaben für Big Data Systeme mit verschiedenen Fallstudien bzw. Datensätzen sowie die Begleitung durch eine Online-Lernplattform unterstützen den Lernerfolg. Die erlernten Inhalte werden im Rahmen einer Projektarbeit praktisch angewendet und vertieft.

## TEILNEHMENDENZAHL

Um eine individuelle Betreuung gewährleisten zu können, ist die Zahl der Teilnehmenden auf maximal 12 beschränkt.

## PRÜFUNG UND ABSCHLUSS

Die Teilnehmenden erhalten eine Teilnahmebescheinigung, wenn mindestens 75% des Kurses besucht werden. Ein Zertifikat der Hochschule Niederrhein wird mit bestandener Prüfungsleistung (schriftliche Dokumentation der Projektarbeit) vergeben.

## PROGRAMM

### I Skalierbare Big Data Architekturen

Selbstlern-einheit	8h	<b>Grundlagen und Begriffe</b> Grundlagen Datenmanagement (Datenmodellierung, Datenbanksysteme, Abfragesprachen); Verteilte Architekturen
Präsenz 1	8h	<b>Big Data Architekturen</b> <ul style="list-style-type: none"><li>• Vertikale und horizontale Skalierbarkeit</li><li>• Grundlagen von verteilten Systemen</li><li>• Cloud-Computing</li><li>• Lambda- und Kappa-Architektur für Big Data</li></ul> <b>Big Data System: Hadoop</b> <ul style="list-style-type: none"><li>• Architektur Hadoop</li><li>• Verteilte Verarbeitung mit Map-Reduce</li><li>• Weitere Big Data Systeme im Hadoop-Ökosystem (z.B. Hive, HBase, YARN)</li><li>• Umsetzung von Datenverarbeitungsprozessen anhand von praktischen Übungen und Fallstudien</li></ul>
Selbstlern-einheit	4h	Weitere Übungen zu Hadoop, Map-Reduce und verwandten Technologien mit praktischen Beispielen; Überblick der Systeme im Hadoop-Ökosystem

### II Big Data Systeme

Präsenz 2	8h	<b>Big Data System: Apache Spark</b> <ul style="list-style-type: none"><li>• Nachteile bei Hadoop und dem Map-Reduce-Programmiermodell</li><li>• Architektur Apache Spark</li><li>• Einführung in die Programmierung einfacher Datenverarbeitungsprozesse in Apache Spark</li><li>• Umsetzung von Datenverarbeitungs- und -analyseprozessen anhand von praktischen Übungen und Fallstudien</li></ul>
Selbstlern-einheit	12h 27h	<ul style="list-style-type: none"><li>• Weitere Übungen zu Apache Spark mit praktischen Beispielen; Rolle von Apache Spark in verteilten Big Data Architekturen</li><li>• Projektarbeit zur Datenverarbeitung und -analyse mit Apache Spark bzw. Hadoop anhand von gegebenen Fallstudien bzw. eigenen Anwendungsfällen aus dem Unternehmen</li></ul>

### III Projektarbeiten & Transfer in die Praxis

Präsenz 3	8h	<b>Vorstellung der Projektarbeiten</b> <ul style="list-style-type: none"><li>• Zusammenfassung der Ergebnisse aus den Projektarbeiten durch den Dozenten</li><li>• Präsentation und Diskussion einzelner Ergebnisse mit den Teilnehmenden und dem Dozenten</li></ul> <b>Praxis-Transfer</b> <ul style="list-style-type: none"><li>• Anwendbarkeit der gelernten Inhalte in der Praxis</li><li>• Präsentation und Diskussion aktueller Trends bei Big Data Systemen</li></ul>
-----------	----	--

Gesamter Zeitaufwand = 75 h, davon Präsenz = 24 h, 3 ECTS

## IHRE ANSPRECHPARTNERIN:

Johanna Tsikouridou-Flesch  
Wissenschaftliche Weiterbildung  
Hochschule Niederrhein  
Reinarzstraße 49 | 47805 Krefeld  
Tel.: 02151 822-1564  
pilotkurse@hs-niederrhein.de

## IHR DOZENT:

Prof. Dr. rer. nat. Christoph Quix  
Wirtschaftsinformatik und Data Science  
Fachbereich Elektrotechnik u. Informatik  
Hochschule Niederrhein



Hochschule Niederrhein  
University of Applied Sciences