

<b>Modulbezeichnung (Kürzel)</b>	<b>Data Science (DASC)</b>	
<b>Modulbezeichnung (eng.)</b>	Data Science	
<b>Semester (Häufigkeit)</b>	5 (jedes Wintersemester)	
<b>ECTS-Punkte (Dauer)</b>	5 (1 Semester)	
<b>Art</b>	Pflichtfach	
<b>Sprache(n)</b>	Deutsch	
<b>Studentische Arbeitsbelastung</b>	60 h Kontaktzeit + 90 h Selbststudium	
<b>Voraussetzungen (laut BPO)</b>		
<b>Empf. Voraussetzungen</b>	Java 1, Java 2, Datenbanken	
<b>Verwendbarkeit</b>	BI, BBTBI, BIPV	
<b>Prüfungsform und -dauer</b>	Klausur 1,5 h oder mündliche Prüfung	
<b>Lehr- und Lernmethoden</b>	Vorlesung, Praktikum	
<b>Modulverantwortliche(r)</b>	T. Schmidt	
<b>Qualifikationsziele</b>		
Die Studierenden kennen die wesentlichen Konzepte in den Bereichen i) Datenintegration und Datenhaltung ii) Datenanalyse und Wissensmanagement sowie iii) Datenvisualisierung und Informationsbereitstellung. Die Studierenden verstehen die Anforderungen von großen Datenmengen (Big Data), kennen grundlegende Konzepte (z.B. MapReduce) und sind mit aktuellen Big-Data Technologien (z.B. Hadoop, Spark) vertraut und können diese auf praktische Problemstellungen anwenden.		
<b>Lehrinhalte</b>		
Vorgestellt werden grundlegende Konzepte und Methoden aus den Data Science Bereichen Maschine Learning/Knowledge Data Discovery in Databases und Big Data die mit praktischen Übungen verdeutlicht werden. Stichworte sind: Bereich KDD/ML:		
<ol style="list-style-type: none"> <li>1. supervised/unsupervised learning</li> <li>2. Algorithmen: clustering (hierarchical, top-down vs. bottom-up, k-means), classification, Decision Trees, Random Forest, Apriori</li> <li>3. Evaluation measures: confusion matrix, ROC, Silhouette, unbalanced classes, challenges &amp; pitfalls.</li> </ol>		
Bereich Big Data:		
<ol style="list-style-type: none"> <li>1. Big Data Collection: cleaning &amp; integration, data platforms &amp; the cloud</li> <li>2. Big Data Storage: Hadoop, modern databases, distributed computing platforms, MapReduce, Spark, NoSQL/NewSQL</li> <li>3. Big Data Systems: Security, Scalability, Visualisation &amp; User Interfaces</li> <li>4. Big Data Analytics: Fast Algorithms, Data Compression, Machine Learning Tools for Big Data Frameworks, Case Studies &amp; Applications (e.g. Medicine, Finance)</li> </ol>		
<b>Literatur</b>		
Freiknecht, Jonas: Big Data in der Praxis: Lösungen mit Hadoop, HBase und Hive. Daten speichern, aufbereiten, visualisieren, Carl Hanser Verlag, 2014		
Karau, Holden: Learning Spark: Lightning-Fast Big Data Analysis, O'Reilly, 2015		
Ester, Martin: Knowledge Discovery in Databases - Techniken und Anwendungen, Springer Verlag, 2000		
<b>Lehrveranstaltungen</b>		
<b>Dozenten/-innen</b>	<b>Titel der Lehrveranstaltung</b>	<b>SWS</b>
T. Schmidt	Data Science	3

