

Modulcode (1.)	Modulbezeichnung (2.)	Zuordnung (3.)
BAI7320	Data Engineering (DE)	
	Studiengang (4.)	Bachelor Angewandte Informatik/ Bachelor Angewandte Informatik DUAL
	Fakultät (5.)	Gebäudetechnik und Informatik

Modulverantwortlich (6.)	Prof. Dr. Nadine Steinmetz
Modulart (7.)	Pflicht Vertiefung WI
Angebotshäufigkeit (8.)	WS
Regelbelegung / Empf. Semester (9.)	BA7
Credits (ECTS) (10.)	5 CP
Leistungsnachweis (11.)	SL - Projekt mit Präsentation
Unterrichtssprache (12.)	Deutsch
Voraussetzungen für dieses Modul (13.)	BAI1050: Datenbanken 1 BAI2050: Datenbanken 2
Modul ist Voraussetzung für (14.)	
Moduldauer (15.)	1 Semester
Notwendige Anmeldung (16.)	-
Verwendbarkeit des Moduls (17.)	Kann von anderen Vertiefungsrichtungen als Wahlmodul belegt werden

Lehrveranstaltung (18.)	Dozent/in (19.)	Art (20.)	Teilnehmer (maximal) (21.)	Anzahl Gruppen (22.)	SWS (23.)	Workload	
						Präsenz (24.)	Selbststudium (25.)
1   Data Engineering	AI17	V/Ü	24	1	4	60	65
Summe					4	60	65
Workload für das Modul (26.)						125	

<b>Qualifikationsziele</b>	<p>(27.)</p> <p>Nach Abschluss dieses Moduls verfügen die Studierenden über grundlegende Kenntnisse in der Auswertung und Analyse umfangreicher Datenbestände. Sie beherrschen Methoden zur Datenaufbereitung und -bereinigung und sind in der Lage, gängige Verfahren aus den Bereichen Data Mining, Machine Learning und OLAP anzuwenden. Durch die Übungen erwerben die Studierenden praktische Erfahrung im Umgang mit gängigen Werkzeugen zur Datenanalyse und -verarbeitung (z. B. Datenbanken, Data Warehouses, interaktive Notebooks). Sie können eigenständig Lösungen zu den gestellten Aufgaben entwickeln und präsentieren, aktiv an fachlichen Diskussionen teilnehmen und aufkommende Fragen kompetent beantworten.</p>
<b>Inhalte</b>	<p>(28.)</p> <ul style="list-style-type: none"> <li>• Arten von Daten</li> <li>• Datenanalyseprozess</li> <li>• Datenvorarbeitung und -bereinigung</li> <li>• Data Warehouse</li> <li>• OLAP</li> <li>• Grundlagen des Data Mining/Machine Learning: ausgewählte Verfahren</li> <li>• Rechtliche Aspekte und Datenschutz</li> <li>• Systeme und Werkzeuge zur Datenanalyse (SQL, Jupyter, Tensorflow)</li> </ul>
<b>Vorleistungen und Modulprüfung</b>	<p>(29.)</p> <ul style="list-style-type: none"> <li>• Vorleistungen: Keine</li> <li>• Leistungsnachweis: Projekt (Gruppenprojekt max. 2 Personen pro Gruppe) mit Präsentation im Vorlesungszeitraum, Gewicht der einzelnen Bestandteile: Präsentation (30%), Dokumentation (30%), Implementierung 40%</li> </ul>
<b>Literatur</b>	<p>(30.)</p> <ul style="list-style-type: none"> <li>• Müller, Lenz: Business Intelligence, Springer, ISBN 978-3-642-35559-2, 2013.</li> <li>• Saake, Sattler, Heuer: Datenbanken - Konzepte und Sprachen, 6. Auflage, mitp-Verlag, 2018.</li> <li>• Van der Plass: Data Science mit Python, mitp-Verlag, 2017.</li> <li>• Kumar, Steinbach, Tan: Introduction to Data Mining, Addison-Wesley, 2005.</li> <li>• Han, Kamber, Pei: Data Mining: Concepts and Techniques, Morgan Kaufmann, 2011.</li> </ul>