

Modulbezeichnung: Data Engineering 2			
Modulkürzel	t.BA.DS.DE2.20HS		
ECTS Credits	4		
Unterrichts- und Prüfungssprache	Deutsch		
Verantwortliche OE	InIT		
Modulverantwortung	Andreas Weiler		
Rechtliche Grundlagen	Die Modulbeschreibung ist neben Rahmenprüfungsordnung und Studienordnung Teil der Rechtsgrundlage. Sie ist verbindlich. Eine in der ersten Unterrichtswoche des Semesters schriftlich festgehaltene und kommunizierte Modulvereinbarung kann die Modulbeschreibung präzisieren. Die Modulvereinbarung ersetzt nicht die Modulbeschreibung.		
Modulprägung	Typ 3a 2 Lektionen Vorlesung pro Semesterwoche und Klasse + 2 Lektionen Praktikum pro Semesterwoche und Halbkasse		
Beschreibung des Moduls	Die verschiedenen Themen des Data Engineering sind existenzielle Bausteine erfolgreicher Datenprodukte und Datenanalyseprojekte. In diesem Modul lernen Sie Voraussetzungen für dessen Einsatz und verschiedene wichtige Methoden des Data Engineering kennen. Es werden theoretische Grundlagen, aber auch praktische Umsetzungen der Methoden gelehrt.		
Inhalte des Moduls	<ul style="list-style-type: none"> Die Digitalisierung von Prozessen und Umgebungen stellt Informatiker vor neue Herausforderungen. Die Softwareentwicklung steht hierbei nicht mehr im Vordergrund, sondern die fachgerechte Verarbeitung und Analyse von verschiedenen Datenarten und Mengen. Hierzu ist es unerlässlich, einen gewissen Grunderfahrungsschatz im Bereich des Data Engineering und der wichtigsten Methoden aus diesem Bereich zu besitzen. Dieses Modul bietet eine praktische Einführung in elementares Data Engineering mittels Methoden zur Verarbeitung von strukturierten und unstrukturierten Daten. Der Fokus liegt hierbei auf einem Überblick und sauberer Methodik. Das Gelernte wird von Teilnehmenden selbständig anhand praktischer Arbeiten vertieft. Storing Structured Data: Relational NoSQL Big Data / Hadoop etc. Transforming Data: ETL Jobs Schedulers Cleaning (Noise removal, Outlier detection, Interpolation) Anonymization Sourcing Data: APIs Web crawling, Web scraping Other Sources Querying Data: Advanced Queries Query Optimization Distributed Queries Ingesting Data: Batch vs. Real-Time Data Streams Queues Die Vorlesung wird begleitet durch praktische Übungen an realen Datensätzen mittels Python und verwandten Tools und Bibliotheken. Etwa die Hälfte des Semesters arbeiten die Teilnehmenden in kleinen Gruppen an individuellen praktischen Arbeiten, die am Ende präsentiert und bewertet werden. 		
Vorkenntnisse	tbd		
Lernziele (Kompetenzen)	Die Studierenden...	Kompetenzen	Taxonomiestufen
	kennen die Perspektiven und Möglichkeiten der aktuellen Forschung und Entwicklung in den genannten Bereichen.	F	K1
	kennen den Unterschied zwischen Data Engineering mit strukturierten, unstrukturierten, Batch und Streaming Daten. Diese werden an Praxisbeispiele mit aktuellen Frameworks angewendet.	M, F	K1, K2, K3
	verstehen die Grundlagen und Besonderheiten des Data Engineering; insbesondere im Gegensatz zu Software Engineering.	F, M	K1, K2
	kennen verschiedene Themen des Data Engineering, insbesondere im Bereich der Transformation und Bereitstellung der Daten, und wenden diese erfolgreich in der Praxis an.	F, M	K1, K2, K3

Modulbezeichnung: Data Engineering 2

Leistungsnachweis	Modulendprüfung	Bewertung	Dauer (Min.)	Gewichtung	Form
	schriftliche Prüfung	Note	90	80	gem. Modulvereinbarung
	Leistungsnachweise während dem Semester	Bewertung	Dauer (Min.)	Gewichtung	Form
	Bericht	Note		20	gem. Modulvereinbarung
Präsenzverpflichtung im Kontaktstudium	Keine				
Lernmaterialien					
Bemerkungen					