

<b>Modulbezeichnung: Data Engineering 1</b>			
<b>Modulkürzel</b>	t.BA.DS.DE1.20HS		
<b>ECTS Credits</b>	4		
<b>Unterrichts- und Prüfungssprache</b>	Deutsch		
<b>Verantwortliche OE</b>	InIT		
<b>Modulverantwortung</b>	Andreas Weiler		
<b>Rechtliche Grundlagen</b>	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.		
<b>Modulprägung</b>	Typ 3a  2 Lektionen Vorlesung pro Semesterwoche und Klasse + 2 Lektionen Praktikum pro Semesterwoche und Halbkasse		
<b>Beschreibung des Moduls</b>	Das Gebiet des "Data Engineering" deckt die wesentlichen Schritte von der Erhebung der Rohdaten bis zur Bereitstellung der validierten, bereinigten Daten für die Nutzung ab. In "Data Engineering 1" werden die Grundlagen des Data Engineerings und die Verarbeitung unstrukturierter Daten thematisiert.		
<b>Inhalte des Moduls</b>	<ul style="list-style-type: none"> <li>• Wir leben in einer Welt, in welcher die Sammlung, Aufbereitung und Nutzung von Daten zentraler denn je ist. Das Gebiet des "Data Engineering" deckt die wesentlichen Schritte von der Erhebung der Rohdaten bis zur Bereitstellung der validierten, bereinigten Daten für die Nutzung ab - Voraussetzung für die folgende Interpretation, das Lernen, oder die visuelle Darstellung. In "Data Engineering 1" werden die Grundlagen des Gebiets und die Verarbeitung von unstrukturierten Daten abgedeckt.</li> <li>• Einführung - Was ist Data Engineering? - Data Engineering im breiteren Kontext der Data Science - Data (Processing) Pipelines - Verschiedene Formen von (Roh-)daten: Big Data, Small Data, Smart Data, ...</li> <li>• Arbeiten mit Daten - Datenformate und Dateiformate (XML, JSON, CSV, ...) - Navigieren in XML/JSON-Daten (XPath, JSONPath) - Tools - Strukturierte vs. unstrukturierte Daten</li> <li>• Verarbeitung verschiedenster Datenarten mit verschiedenen NoSQL Lösungen</li> <li>• Grundlagen des Information Retrievals (IR)</li> </ul>		
<b>Vorkenntnisse</b>	<a href="https://gpmpublic.zhaw.ch/GPMDocProdDPublic/2_ Studium/2_02_ Grundlagen_ Studium/T_C L_ Modulauspraegungen_ SM2025.pdf">https://gpmpublic.zhaw.ch/GPMDocProdDPublic/2_ Studium/2_02_ Grundlagen_ Studium/T_C L_ Modulauspraegungen_ SM2025.pdf</a>		
<b>Lernziele (Kompetenzen)</b>	<b>Die Studierenden...</b>	<b>Kompetenzen</b>	<b>Taxonomiestufen</b>
	wissen, wie unstrukturierte Daten verarbeitet und Informationen aus Ihnen gewonnen werden kann.	F	K2
	kennen die Grundlagen des Data Engineering	F	K1
	können mittels NoSQL Techniken verschiedenste Arten von Daten verarbeiten, abfragen und verfügbar machen.	F	K3
	verstehen, wie mittels Data Pipelines die Rohdaten für die spätere Nutzung erhoben, aufbereitet und bereinigt werden, und können solche Pipelines designen und implementieren	F	K2, K3

## Modulbezeichnung: Data Engineering 1

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