Valid from 2024.HS

Module description: Data Engineering 1

Module Code	t.BA.DS.DE1.20HS
ECTS Credits	4
Language of Instruction/Examination	German
Organizational Unit	InIT
Module Coordinator	Andreas Weiler
Legal Framework	The module description is part of the legal basis in addition to the general academic regulations. It is binding. During the first week of the semester a written and communicated supplement can specify the module description in more detail.
Module Characteristic	Type 3a 2 lecture lessons per semester week and class+ 2 lab bi-weekly lessons per semester and half-class
Module Description	The field of "Data Engineering" covers the crucial steps from acquisition of the raw data to making the validated, cleaned data available for exploitation. The "Data Engineering 1" module discusses the basics of this field and the handling of unstructured data.
Module Content	• We live in a world in which the collection, transformation and exploitation of data is more central than ever. The field of "Data Engineering" covers the crucial steps from acquisition of raw data to making the validated, cleaned data available for exploitation - such as interpretation, learning and visual rendering. The module "Data Engineering 1" discusses the basics of the field and the handling of unstructured data. • Introduction - What is Data Engineering? - Data Engineering in the broader context of Data Science - Data (Processing) Pipelines - Different forms of data: Big Data, Small Data, Smart Data, .. • Working with data - Data formats and file formats (XML, JSON, CSV, ...) - Navigating XML/JSON data (XPath, JSONPath) - Tools - Structured vs. unstructured data • Handling of different data types and using different NoSQL solutions • Foundations of Information Retrieval (IR)
Prerequisite Knowledge	https://gpmpublic.zhaw.ch/GPMDocProdDPublic/2_Studium/2_02_Grundlagen_Studium/T_CL_Modulauspraegungen_SM2025.pdf

Learning Objectives (Competences)

Students...	Competencies	Taxonomies
know how unstructured data is processed and how information extraction can be achieved.	F	K2
know the basics of Data Engineering	F	K1
can use NoSQL technologies to process, query and access miscellaneous types of data.	F	K3
understand how data pipelines are used for acquiring, transforming and cleaning raw data, and you know how to design and implement such pipelines	F	K2, K3

Performance Assessment

End-of-module exam	Assessment	Length (min.)	Weighting	Form
written exam	Grade	90	80	acc. to module agreement

Performance assessment during the semester	Assessment	Length (min.)	Weighting	Form
report	Grade		20	acc. to module agreement

Module description: Data Engineering 1

Classroom Attendance Requirement	None
Learning material	
Comments	