

| Module description: Data Engineering 2 | | | |
|---|---|---------------------|-------------------|
| Module Code | t.BA.DS.DE2.20HS | | |
| 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 | Data Engineering topics are essential components of successful data products and data projects. Students learn the requirements for running successful data engineering pipelines, the key methods, and both the theoretical foundations and practical implementation of different methods and applications. | | |
| Module Content | <ul style="list-style-type: none"> • The digitalization of processes and environments is difficult challenge for computer scientists. Software development is hereby not the primary problem, rather the professional processing and analysis of different datatypes and volumes. For this purpose it is essential to have a certain fundamental experience in the area of data engineering. • 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 • The lecture is accompanied by practical assignments containing of implementations in python and related tools and libraries with real-world datasets. About the half of the semester the students work in small groups on individual projects, which are presented at the end of the semester. | | |
| Prerequisite Knowledge | tbd | | |
| Learning Objectives (Competences) | Students... | Competencies | Taxonomies |
| | know perspectives and opportunities of current research and development in the domain of data engineering. | F | K1 |
| | know the difference between data engineering of structured, unstructured, batch and stream processing data. | M, F | K1, K2, K3 |
| | understand the fundamentals and specialities of data engineering, especially in contrast to Software development projects. | F, M | K1, K2 |
| | know different topics of data engineering, especially in the domain of transformation and delivery of data. | F, M | K1, K2, K3 |

Module description: Data Engineering 2

| | | | | | | |
|---|---|-------------------|----------------------|----------------------|--------------------------|--------------------------|
| 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 |
| Classroom Attendance Requirement | None | | | | | |
| Learning material | | | | | | |
| Comments | | | | | | |