

Program of Studies:	Master Program Bioinformatics
Name of the module:	Database Systems
Abbreviation:	I-M-3
Subtitle:	Core Lecture
Modules:	Lecture: 4 h (weekly) Tutorial: 2 h (weekly)
Semester:	1 st -3 rd semester/at least once every two years
Responsible lecturer:	Prof. Dr. Jens Dittrich
Lecturer:	Prof. Dr. Jens Dittrich
Language:	English
Level of the unit/ Mandatory or not:	Graduate course / mandatory elective
Total workload:	270 h = 90 h of classes and 180 h private study
Credits:	9
Entrance requirements:	<p>For graduate students:</p> <ul style="list-style-type: none"> - motivation for databases and database management systems - the relational data model - relational query languages, particularly relational algebra and SQL - solid programming skills in Java and/or C++ - undergrad courses in algorithms and data structures, concurrent programming
Aims/Competences to be developed:	<p>Database systems are the backbone of most modern information systems and a core technology without which today's economy – as well as many other aspects of our lives – would be impossible in their present forms. The course teaches the architectural and algorithmic foundations of modern database management systems (DBMS), focussing on database systems internals rather than applications. Emphasis is made on robust and time-tested techniques that have led databases to be considered a mature technology and one of the greatest success stories in computer science. At the same time, opportunities for exciting research in this field will be pointed out.</p> <p>In the exercise part of the course, important components of a DBMS will be treated and where possible implemented and their performance evaluated. The goal this is to work with the techniques introduced in the lecture and to understand them and their practical implications to a depth that would not be attainable by purely theoretical study.</p>

Content:	<p>The course "Database Systems" will introduce students to the internal workings of a DBMS, in particular:</p> <ul style="list-style-type: none"> - storage media (disk, flash, main memory, caches, and any other future storage medium) - data managing architectures (DBMS, streams, file systems, clouds, appliances) - storage management (DB-file systems, raw devices, write-strategies, differential files, buffer management) - data layouts (horizontal and vertical partitioning, columns, hybrid mappings, compression, defragmentation) - indexing (one- and multidimensional, tree-structured, hash-, partition-based, bulk-loading and external sorting, differential indexing, read- and write-optimized indexing, data warehouse indexing, main-memory indexes, sparse and dense, direct and indirect, clustered and unclustered, main memory versus disk and/or flash-based) - processing models (operator model, pipeline models, push and pull, block-based iteration, vectorization, query compilation) processing implementations (join algorithms for relational data, grouping and early aggregation, filtering) - query processing (scanning, plan computation, SIMD) - query optimization (query rewrite, cost models, cost-based optimization, join order, join graph, plan enumeration) - data recovery (single versus multiple instance, logging, ARIES) - parallelization of data and queries (horizontal and vertical partitioning, shared-nothing, replication, distributed query processing, NoSQL, MapReduce, Hadoop and/or similar and/or future systems) - read-optimized system concepts (search engines, data warehouses, OLAP) - write-optimized system concepts (OLTP, streaming data) - management of geographical data (GIS, google maps and similar tools) - main-memory techniques
Assessment/Exams:	<ul style="list-style-type: none"> - Passing a two-hour written exam at the end of the semester - Successful demonstration of programming project (teams of up to three students are allowed); the project may be integrated to be part of the weekly assignments <p>Grades are based on written exam; 50% in weekly assignments (in paper and additionally paper or electronic quizzes) must be passed to participate in the final and repetition exams. A repetition exam takes place during the last two weeks before the start of lectures in the following semester.</p>
Grade:	Will be determined based on project, midterm and best of endterm and reexam.
Literature:	Will be announced before the start of the course on the course page on the Internet.