ACLS

Databases
and
Data Architecture Systems

Spring 2020

About me

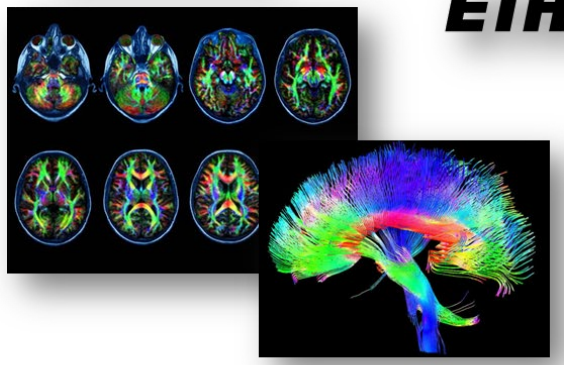
Dr. Robert Vorburger

Head of Research Group
Knowledge Engineering
IAS, LSFM, ZHAW

- MSc ITET, ETH
- PhD IBT, ETH
- Helveting Engineering AG
- IAS, ZHAW

UniversitätsSpital Zürich

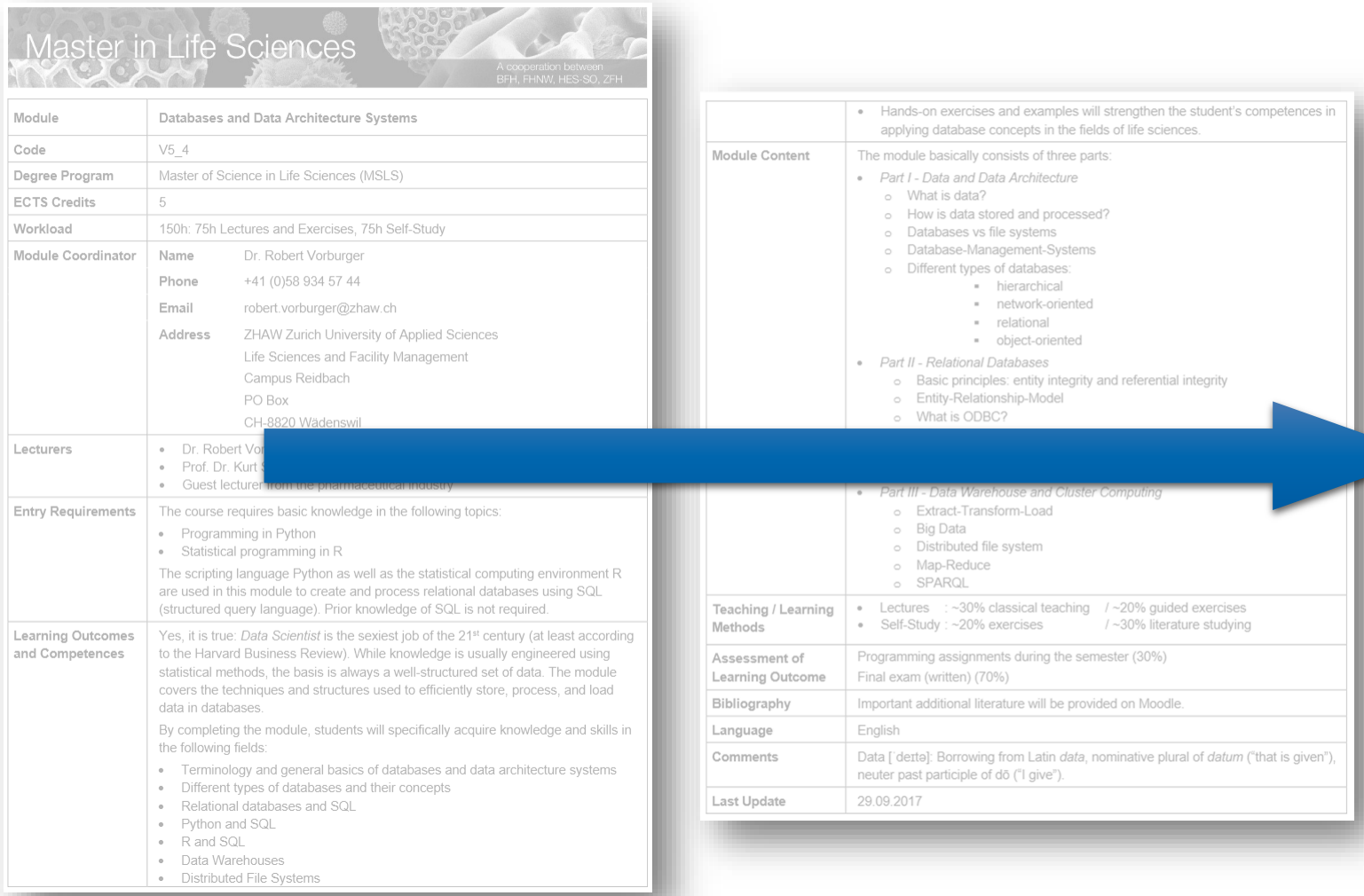
ETH zürich



Columbia University

HELVETING
Pure Passion.

About this Course



Master in Life Sciences

A cooperation between BFH, FHNW, HES-SO, ZFH

Module	**Databases and Data Architecture Systems**
Code	V5_4
Degree Program	Master of Science in Life Sciences (MSLS)
ECTS Credits	5
Workload	150h: 75h Lectures and Exercises, 75h Self-Study
Module Coordinator	**Name** Dr. Robert Vorburger **Phone** +41 (0)58 934 57 44 **Email** robert.vorburger@zhaw.ch **Address** ZHAW Zurich University of Applied Sciences Life Sciences and Facility Management Campus Reidbach PO Box CH-8820 Wädenswil
Lecturers	• Dr. Robert Vo • Prof. Dr. Kurt S • Guest lecturer from the pharmaceutical industry
Entry Requirements	The course requires basic knowledge in the following topics: • Programming in Python • Statistical programming in R The scripting language Python as well as the statistical computing environment R are used in this module to create and process relational databases using SQL (structured query language). Prior knowledge of SQL is not required.
Learning Outcomes and Competences	Yes, it is true: *Data Scientist* is the sexiest job of the 21st century (at least according to the Harvard Business Review). While knowledge is usually engineered using statistical methods, the basis is always a well-structured set of data. The module covers the techniques and structures used to efficiently store, process, and load data in databases. By completing the module, students will specifically acquire knowledge and skills in the following fields: • Terminology and general basics of databases and data architecture systems • Different types of databases and their concepts • Relational databases and SQL • Python and SQL • R and SQL • Data Warehouses • Distributed File Systems

	• Hands-on exercises and examples will strengthen the student's competences in applying database concepts in the fields of life sciences.

Module Content	The module basically consists of three parts:
	• *Part I - Data and Data Architecture*
	  ○ What is data?
	  ○ How is data stored and processed?
	  ○ Databases vs file systems
	  ○ Database-Management-Systems
	  ○ Different types of databases:
	  ▪ hierarchical
	  ▪ network-oriented
	  ▪ relational
	  ▪ object-oriented
	• *Part II - Relational Databases*
	  ○ Basic principles: entity integrity and referential integrity
	  ○ Entity-Relationship-Model
	  ○ What is ODBC?
	• *Part III - Data Warehouse and Cluster Computing*
	  ○ Extract-Transform-Load
	  ○ Big Data
	  ○ Distributed file system
	  ○ Map-Reduce
	  ○ SPARQL

Teaching / Learning Methods	• Lectures : ~30% classical teaching / ~20% guided exercises • Self-Study : ~20% exercises / ~30% literature studying
Assessment of Learning Outcome	Programming assignments during the semester (30%) Final exam (written) (70%)
Bibliography	Important additional literature will be provided on Moodle.
Language	English
Comments	Data [ˈdeɪtə]: Borrowing from Latin *data*, nominative plural of *datum* ("that is given"), neuter past participle of dō ("I give").
Last Update	29.09.2017

Know how to…

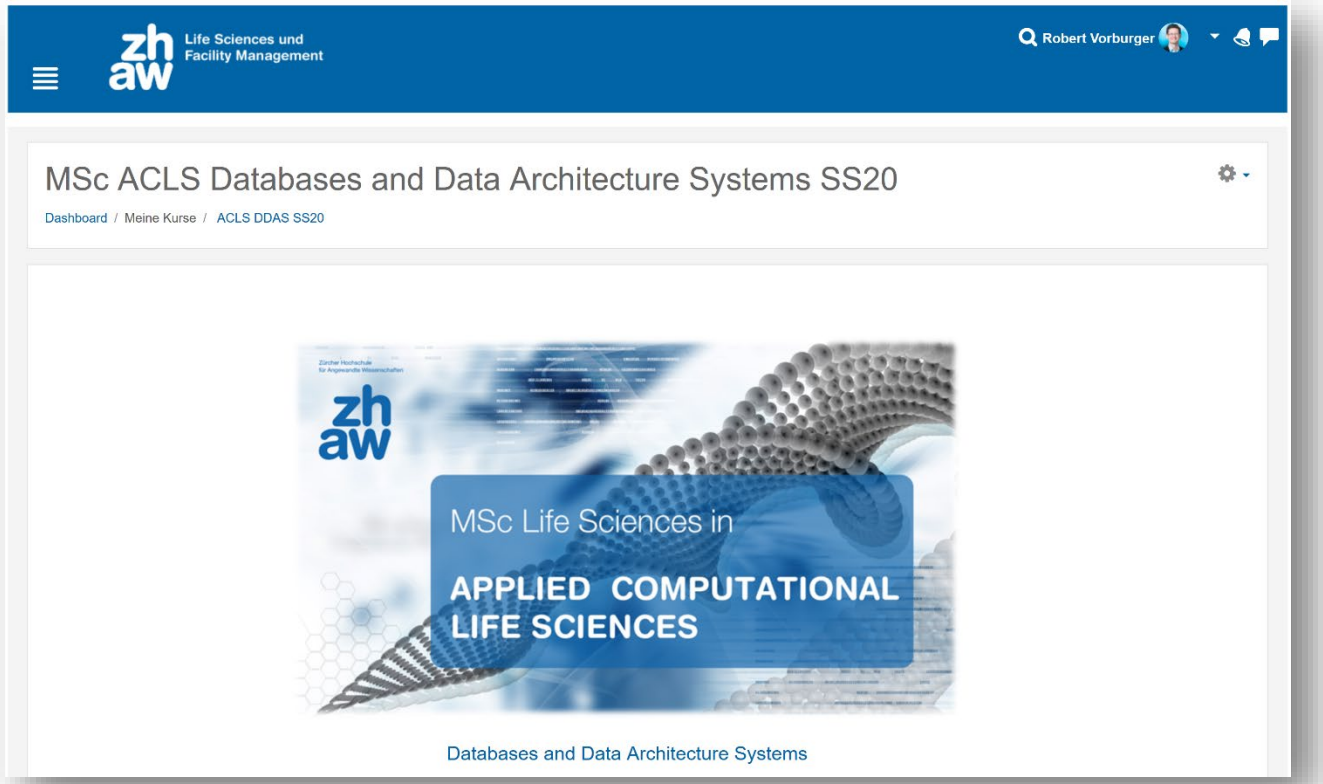
- …load data from a database
- …store data in a database

Schedule

Calendar Week	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Module Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Date	17.02.	24.02.	02.03.	09.03.	16.03.	23.03.	30.03.	06.04.	13.04.	20.04.	27.04.	04.05.	11.05.	18.05.
Topic		SQL	SQL								NoSQL			

- Introduction & RDB
- Structured Query Language
- Database & Python
- Database & R
- Data Warehouse
- Not only SQL
- Graph Database
- Special

Moodle



➡ moodle.zhaw.ch

Course

ACLS DDAS SS20

Grading

Coding Project 30%

Written Exam 70%

Calendar Week	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Module Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Date	17.02.	24.02.	02.03.	09.03.	16.03.	23.03.	30.03.	06.04.	13.04.	20.04.	27.04.	04.05.	11.05.	18.05.
Topic		SQL	SQL				R				NoSQL			

February March April May

15.06.2020
10:00-11:30

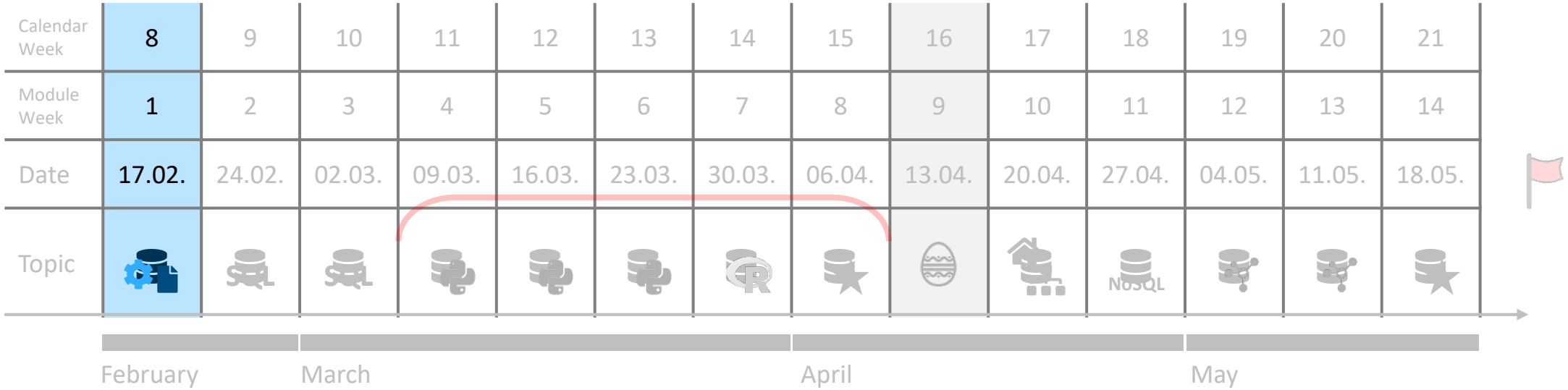
ACLS

Databases

SS 2020 – Week 1

February 17

Schedule



Introduction & RDB

Structured Query Language

Database & Python

Database & R

Data Warehouse

Not only SQL

Graph Database

Special

Learning Objectives

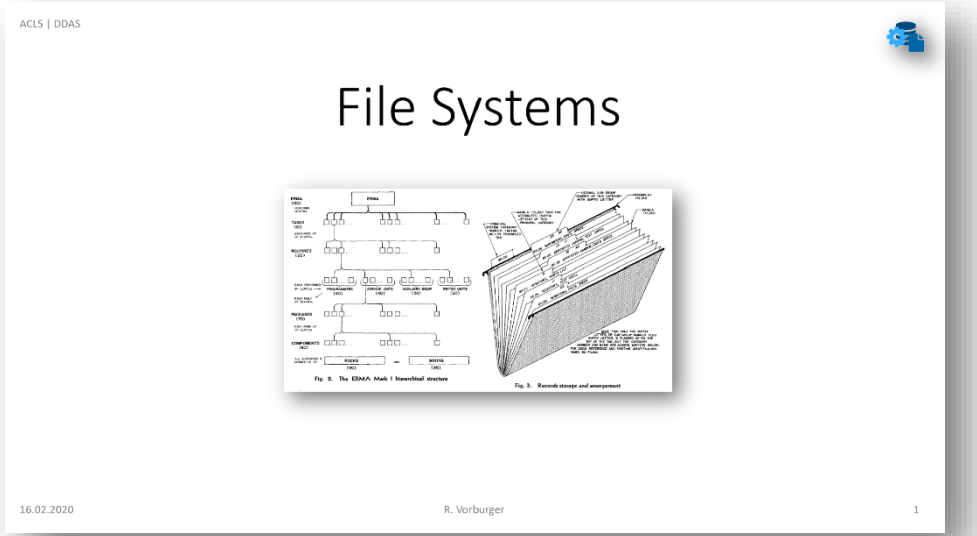
- You know why relational databases are called relational
- You know what relational integrity is
- You are able to define normalized tables for a relational database
- You are able to design the Entity-Relationship-Model (ERM) for a relational database

Content

• Introduction		20’
• File Systems	🎓	20’
• Databases	🎓	45’
• Relational Databases	🎓	40’
• Relational Databases	🔧	30’
• Entity-Relationship-Model	🎓	40’
• Entity-Relationship-Model	🔧	30’

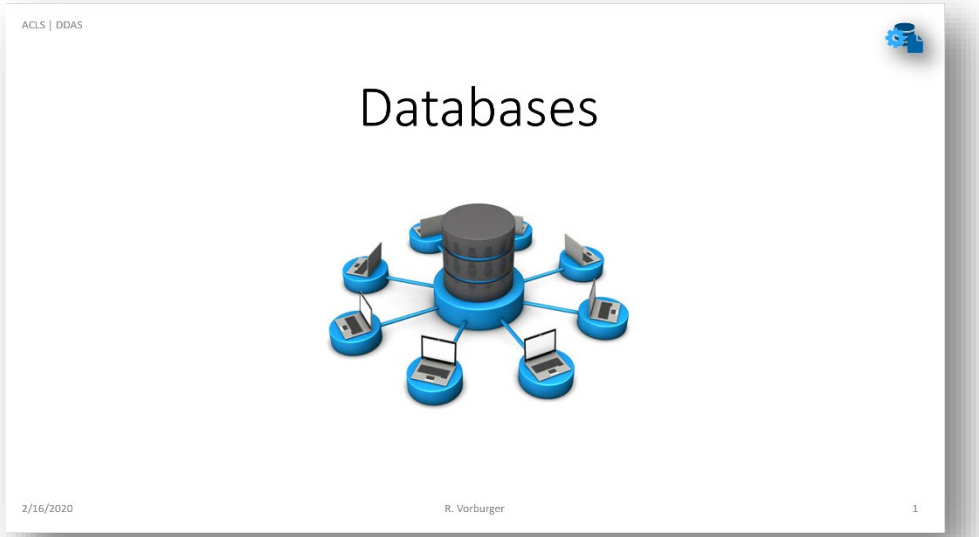
Content

• Introduction		20’
• File Systems	🎓	20’
• Databases	🎓	45’
• Relational Databases	🎓	40’
• Relational Databases	🔧	30’
• Entity-Relationship-Model	🎓	40’
• Entity-Relationship-Model	🔧	30’



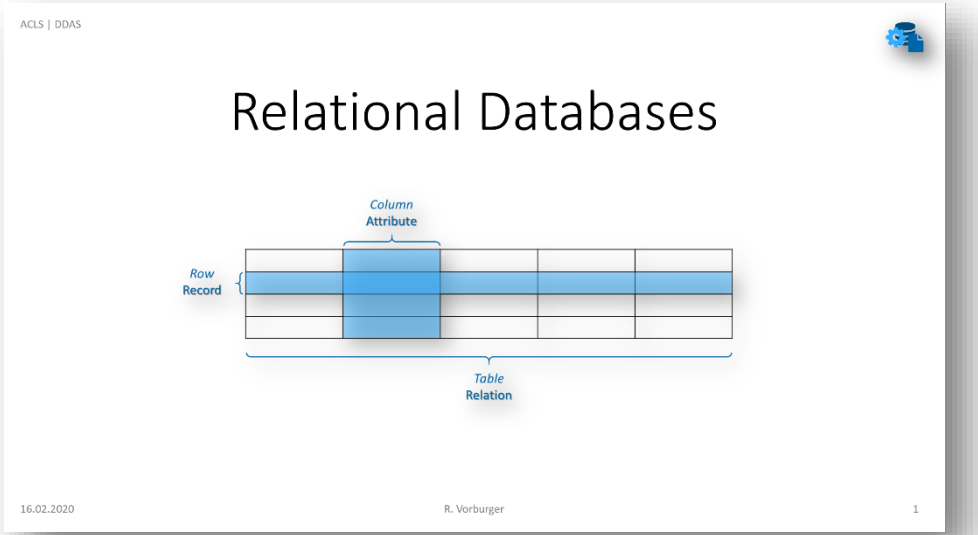
Content

• Introduction		20’
• File Systems	🎓	20’
• **Databases**	🎓	45’
• Relational Databases	🎓	40’
• Relational Databases	🔧	30’
• Entity-Relationship-Model	🎓	40’
• Entity-Relationship-Model	🔧	30’



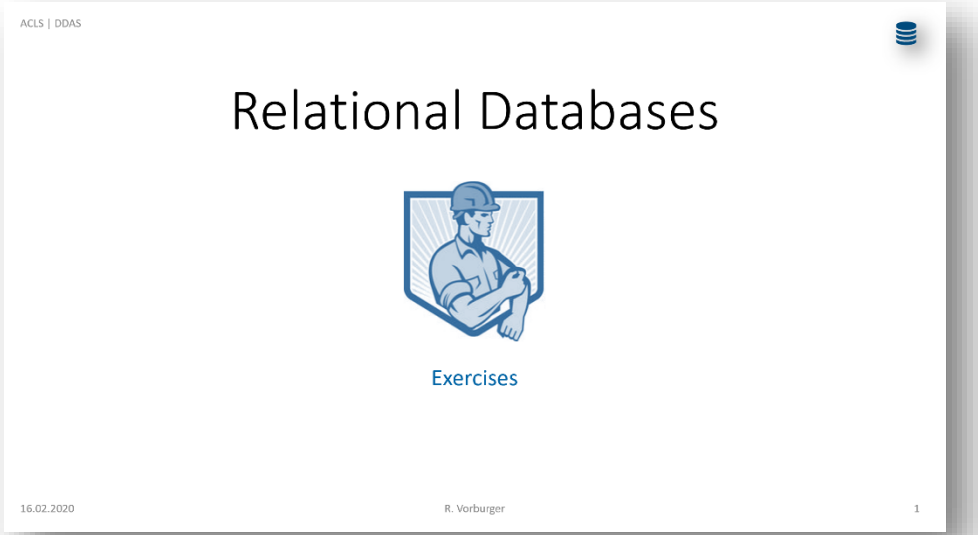
Content

• Introduction		20’
• File Systems		20’
• Databases		45’
• Relational Databases		40’
• Relational Databases		30’
• Entity-Relationship-Model		40’
• Entity-Relationship-Model		30’



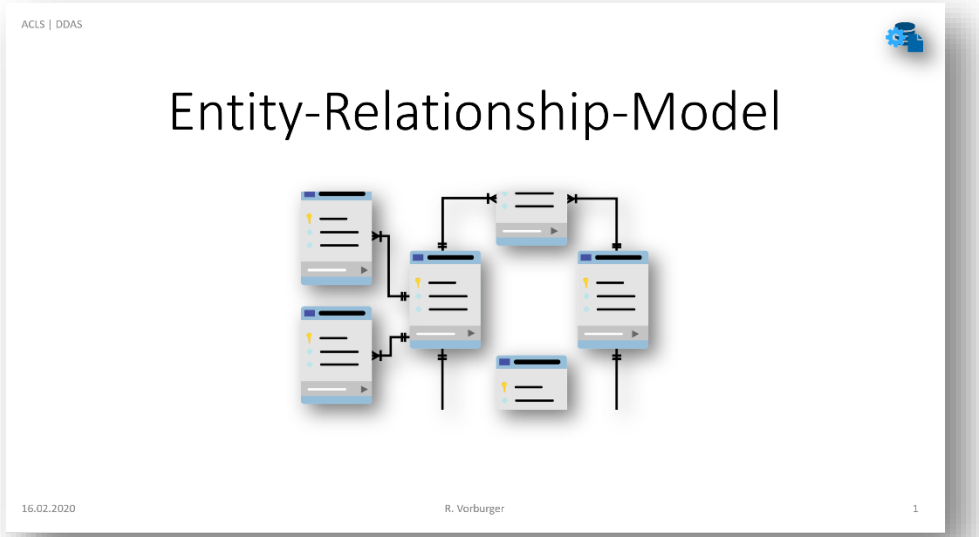
Content

• Introduction		20’
• File Systems		20’
• Databases		45’
• Relational Databases		40’
• Relational Databases		30’
• Entity-Relationship-Model		40’
• Entity-Relationship-Model		30’



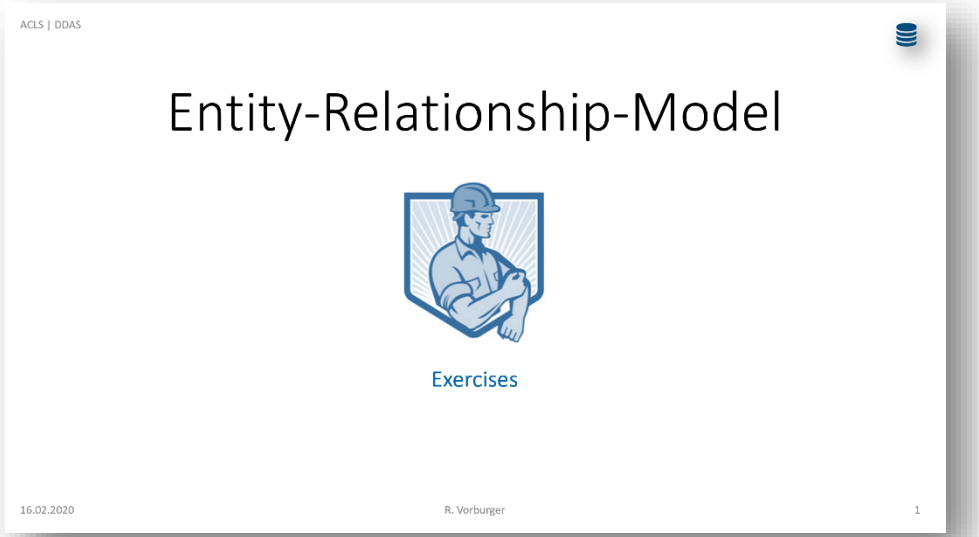
Content

• Introduction		20’
• File Systems	🎓	20’
• Databases	🎓	45’
• Relational Databases	🎓	40’
• Relational Databases	🔧	30’
• **Entity-Relationship-Model**	🎓	40’
• Entity-Relationship-Model	🔧	30’



Content

• Introduction		20’
• File Systems	🎓	20’
• Databases	🎓	45’
• Relational Databases	🎓	40’
• Relational Databases	🔧	30’
• Entity-Relationship-Model	🎓	40’
• Entity-Relationship-Model	🔧	30’



Learning Objectives

r = { (0,0,0), (0,1,-1), (0,2,-2),
(1,0,1), (1,1,0), (1,2,-1),
(2,0,2), (2,1,1), (2,2,0) }

subtraction

minuend	subtrahend	difference
0	0	0
0	1	-1
0	2	-2
1	0	1
1	1	0
1	2	-1
2	0	2
2	1	1
2	2	0

- You know why relational databases are called relational
- You know what relational integrity is
- You are able to define normalized tables for a relational database
- You are able to design the Entity-Relationship-Model (ERM) for a relational database

employee

EmployeeNr	Department	Name	Age
MN0345	MN	John Smith	52
SE2376	SE	Peter Johnson	29
SE8568	SE	Alice Winter	32
MN3785	MN	Mary Jones	24
MN9448	MN	Peter McAlister	47

project

ProjectNr	PM	Budget
P870	MN3785	50'000
P348	MN0345	120'000
P101	MN0345	1'000

PK is used as FK in another table.
Record cannot be deleted.