

# UV Distributed Information Management

## Introduction

Martin Schäler, [martin.schaeler@plus.ac.at](mailto:martin.schaeler@plus.ac.at)

special thanks to Daniel Kocher

Department of Computer Science  
University of Salzburg

# Agenda

---

# Today's Agenda

1. Introduction
2. Logistics
3. Outlook
4. Discussion

# Who Am I?

1. **Martin Schäler** (Associate Prof. for Data-Intensive Systems)
2. At PLUS since 2020
3. Before: Univ. Karlsruhe, Univ. Magdeburg

## Contact (in this order):

1. **Lecture:** Wednesdays, 11:00 - 13:00 c.t.
2. **Email:** martin.schaeeler@plus.ac.at.
3. Try to avoid Blackboards message system.

Please **interrupt me** if

- you have **troubles understanding** what I am talking about,
- you have a **question** related to the current topic, or

# Who are you?

- Name
- Background knowledge: Databases and Programming
- What (new) skills would you like to have at after passing this course?
- Whatever you deem important?

## Why You Should Take This Course

The amount of **information** is growing continuously and **needs to be managed**.

**Many systems** for different scenarios have been developed.

You must be able to **choose the proper system** for your use case.

## Topic Outline

---

# Distributed Information Management



Data Management

Data Processing



# Distributed Information Management



Data Management

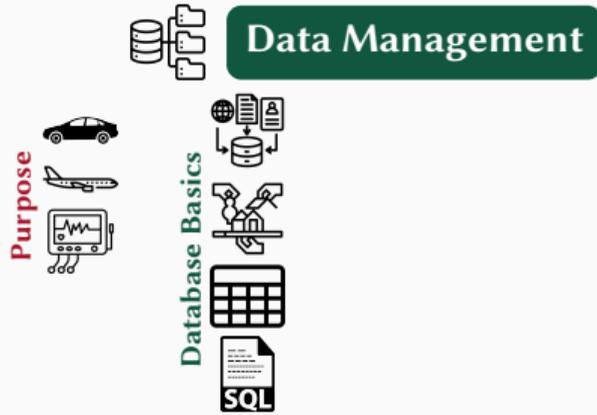
Data Processing



Purpose



# Distributed Information Management



# Distributed Information Management

## Data Management

Purpose



Database Basics



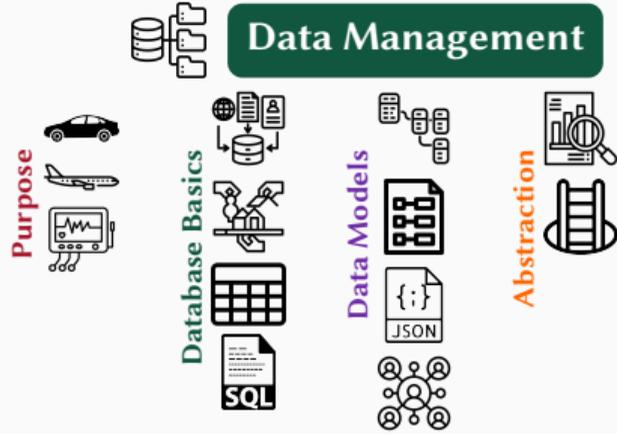
Data Models



## Data Processing



# Distributed Information Management



## Data Processing



# Distributed Information Management

Purpose



Database Basics



Data Models



Abstraction



## Data Management

## Data Processing



Workloads



# Distributed Information Management

## Data Management

Purpose



Database Basics



Data Models



Abstraction



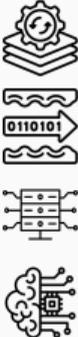
Challenges



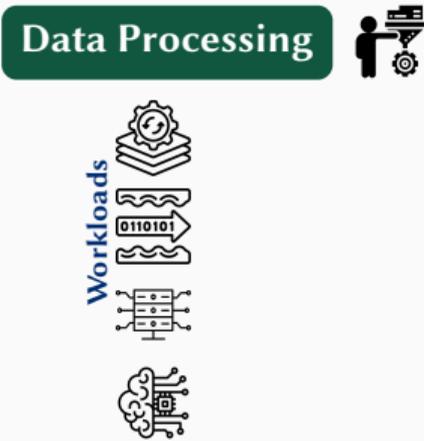
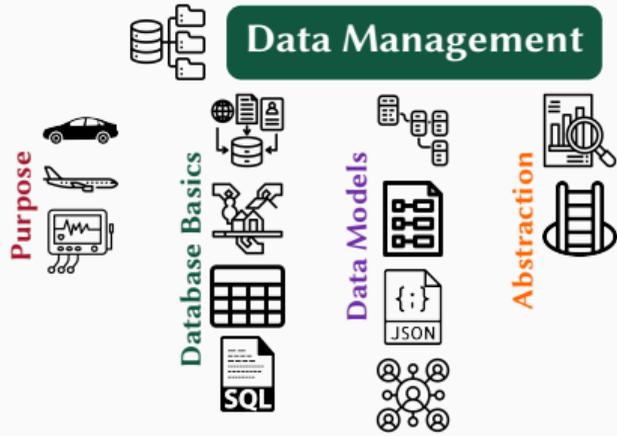
## Data Processing



Workloads



# Distributed Information Management



# Distributed Information Management



**Figure 1:** Generated using Stable Diffusion WebUI:  
<https://github.com/AUTOMATIC1111/stable-diffusion-webui>

## Complementary Courses

We will cover some basics of database management systems, but **not the internals in detail**. For internals, we refer to other courses at the Department of Computer Science.

### Undergraduate Courses:

- Databases 1 (VO+PS; summer semester)
- Databases 2 (VO+PS; winter semester)
- Database Tuning (VO+PS; summer semester)
- Data management Beyond Relations (VU; summer semester)

### Graduate Courses:

- Advanced Databases (VO+PS; winter semester)
- Non-Standard Database Systems (VO+PS; summer semester)
- Similarity Search in Large Databases (VO+PS; winter semester)

## Course Objectives

- Develop an **intuition** for database systems and their **core principles**.
- Be able to **properly use the terminology** that is commonly used in these systems.
- **Tackle simple problems** using these systems.
- Learn to **choose a system** for a given problem/application context.
- **Hands-on experience** with selected systems.
- Working in a **team** (organize yourself as a team).
- Learn about the **challenges and limitations** of these systems.

VU “Grundlagen Informatik und Systeme” (512.023)

⇒ **Core principles** of computer science and systems.

UE “Einführung in Programmieren mit Python” (512.024)

⇒ **Basic programming skills** (e.g., in Python3).

# Logistics



# Where to Find Information

## Website<sup>1</sup>:

- Assignments: schedule, meetings, and late submission policy.
- Detailed grading scheme.
- Learning material: slides, hands-on instructions, and supplementary material.
- Midterms: schedule, topics, and modes.

## Blackboard<sup>2</sup>:

- Announcements (important announcements also via email).
- Assignments: announcement and submission.

---

<sup>1</sup>Course website: <https://dbresearch.uni-salzburg.at/teaching/2025ss/dim/>

<sup>2</sup>PLUS Blackboard: <https://elearn.sbg.ac.at>

## General Information:

- Only **one group**: **Wednesdays, 11:00 - 13:30h c.t.**, T 02.
- Lecture will be in **in person**
- However, **attendance** is **obligatory** and will be checked (*“prüfungsimmanent”*).
- **Theoretical background** for the **assignments**.
- Covers all **relevant topics** for the **exams**.

## General Information:

- **Groups of 2-3** students (enrollment via Blackboard<sup>3</sup>).
- **Practical assignments** related to the topics covered in class.
- Submission via **Blackboard** (4 weeks per assignment; max. 5 weeks).
- **3 assignments, each contributing 20%** to your grade.
- Each group presents the results to me (After-Assignment Meeting)

---

<sup>3</sup>If there is any problem in the Blackboard course, please notify me as soon as possible.

# Assignment 0 – Get Familiar with the Assignment Workflow

## Tasks:

1. Find yourself a **team of 2-3 students**.
2. **Enroll your team** in the Blackboard course.
3. Play around with **Debian Linux** (optional).
4. Prepare a live Demo for the After-Assignment Meeting.
5. Submit the report/questionnaire to Blackboard

**Deadline:** March 13, 2024, 11:55 pm (aka 23:55) CET

**Optional meeting** (for 5% bonus) and **no late submission** for this test assignment.

Students without a team after the deadline are **randomly assigned to new groups**.

## General Information:

- Exam is **open book** and will be held **in person**
- Contributes **40%** to your grade.
- Exam will last **at most 90 minutes**.
- Scheduled Appointment June, 25th.

The **exam date** and **time** is the **same for all students**.

# Schedule Overview

	March				April				May				June					
Date	05	12	19	26	02	09	16	23	30	07	14	21	28	04	11	18	25	
VO	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Assignments	Ass. 0		👥				Ass. 2				👥							
			Ass. 1				👥				Ass. 3				👥			
Exams																	📄	📄

CW Calender Week

✓ Regular lecture (in person / hybrid)

✗ No lecture

▶ Remote lecture or video recordings to watch

👥 After-Assignment Meetings

📄 Midterm / Final Exam

📄 Question & Answer session

# Grading Scheme

Assignments		Exams		Total
Assignment 0	5% (Bonus)			
Assignment 1	20%			
Assignment 2	20%			
Assignment 3	20%	Final exam	40%	
	<b>60%</b>		<b>40%</b>	<b>100%</b>

# Grading Scheme

<b>Overall Points</b>	<b>Grade</b>
$\geq 88.75\%$	1 – “Sehr Gut”
[77.5%, 88.75%)	2 – “Gut”
[66.25%, 77.5%)	3 – “Befriedigend”
[55%, 66.25%)	4 – “Genügend”
$< 55\%$	5 – “Nicht Genügend”

**Q&A**

## In-class Exercise:

1. Find yourself a group or work alone – **as you prefer.**
2. The next slide lists **28 + 1 terms.**
3. Try to briefly **describe each term** in the **context of computer science.**

**Time:** 30 minutes

Afterwards, we discuss these **terms in the plenum.**

# Distributed Information Management

Throughput

Concurrency

Processor / CPU

Efficiency

Trade-off

HDD

Database system

Transaction

Declarative

Address

Instruction(s)

Parallelism

Main memory (RAM)

Performance

Virtual machine

SSD

Query

Scalability

Terminal

Time

Space

Cache

Transparency

Redundancy

Memory hierarchy

Index

Physical level

Compiler

## Icons (pictograms made by)

---

mynamepong

Freepik

juicy\_fish

Uniconlabs

Smashicons

Eucalyp

Elzicon

Skyclick

Muhammad Atif

surang

ultimatearm

imaginationlol

Pixel perfect

---

from <https://www.flaticon.com>