

# Vertiefungsrichtung **Datenbanksysteme**

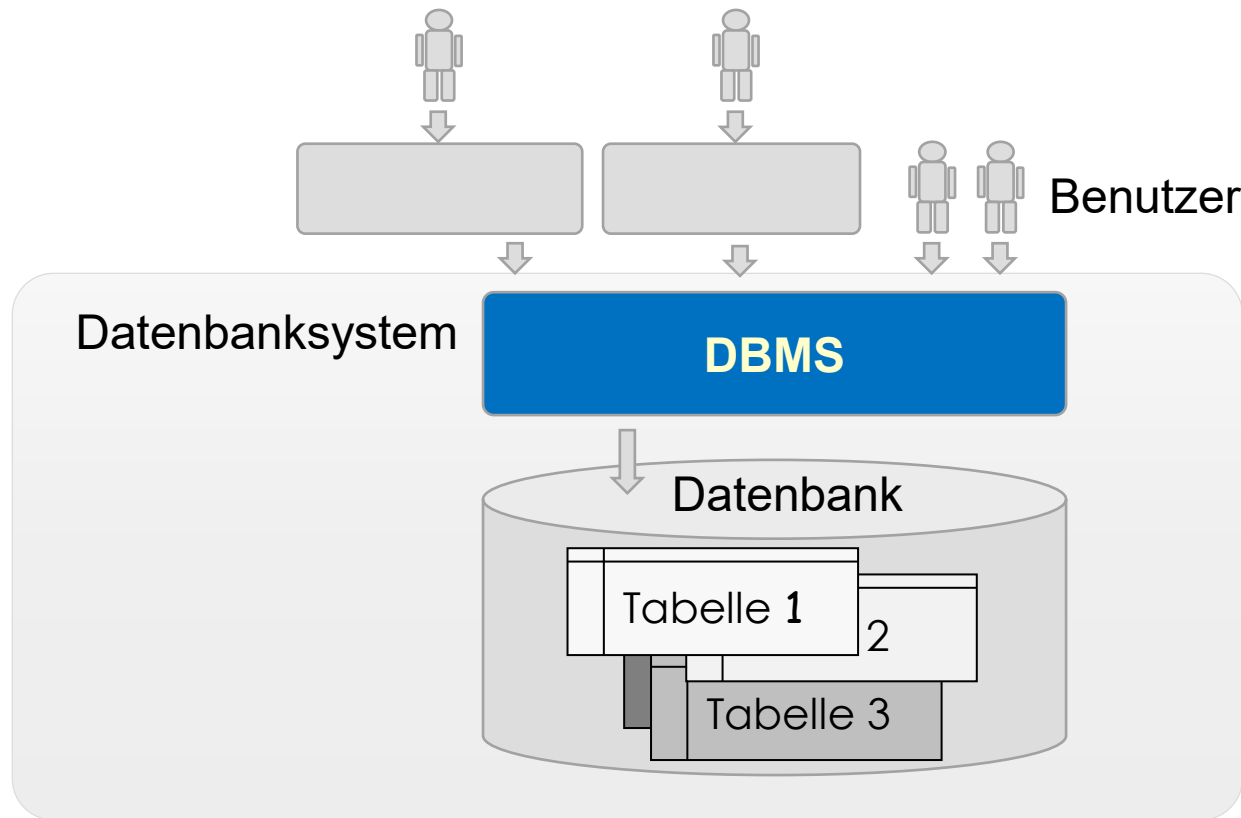
## Orientierungsvorlesung

Richard Lenz

Sommersemester 2023



## Techniken zum **fehlertoleranten** und **effizienten** Zugriff auf **gemeinsam benutzte** Ressourcen (Daten)



*Zur Erinnerung aus den  
Einführungsvorlesungen*

*Ein wichtiges Ziel:  
**Datenunabhängigkeit***



## Pflichtvorlesungen (bei Studienbeginn vor WS 2022):

### 1. Semester: Konzeptionelle Modellierung

- Datenbankentwurf
- E/R-Modell
- Relationales Datenmodell
- UML / XML / Ontologien

*Künftig nur noch im  
Wintersemester  
letztmalig WS 2023/2024*

### 5. Semester: Implementierung von Datenbanksystemen

- Funktionsweise von Datenbanksystemen
- Effizienter Datenbankzugriff: Zugriffspfade
- Fehlertoleranz und Konsistenzsicherung: Transaktionen

*Wintersemester*

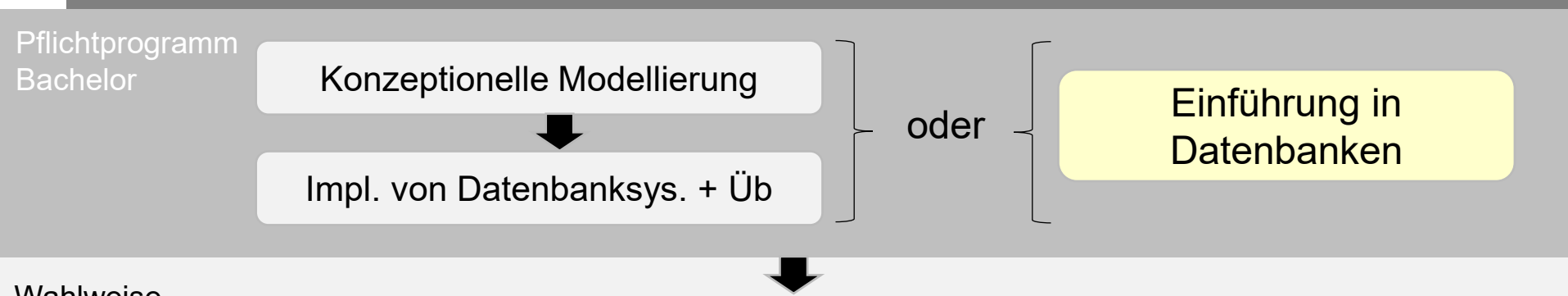


## Pflichtvorlesungen (bei Studienbeginn ab WS 2022):

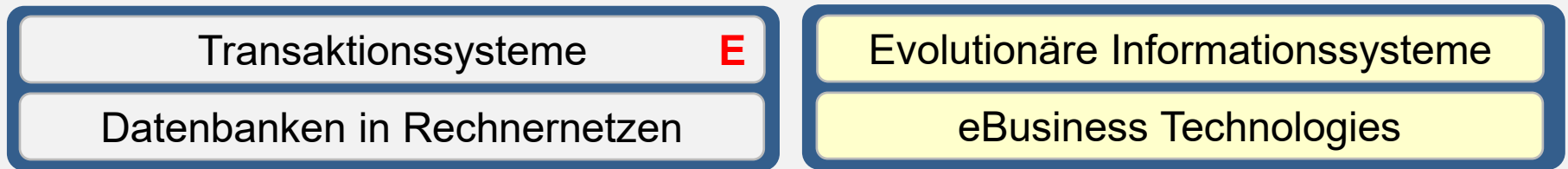
4. Semester: Einführung in Datenbanken (EDB) **7,5 ECTS**

*Künftig in jedem  
Sommersemester*

- Datenbankentwurf
- E/R-Modell
- Relationales Datenmodell
- Funktionsweise von Datenbanksystemen
- Effizienter Datenbankzugriff: Zugriffspfade
- Fehlertoleranz und Konsistenzsicherung: Transaktionen



Wahlweise  
Bachelor oder Master



## Legende

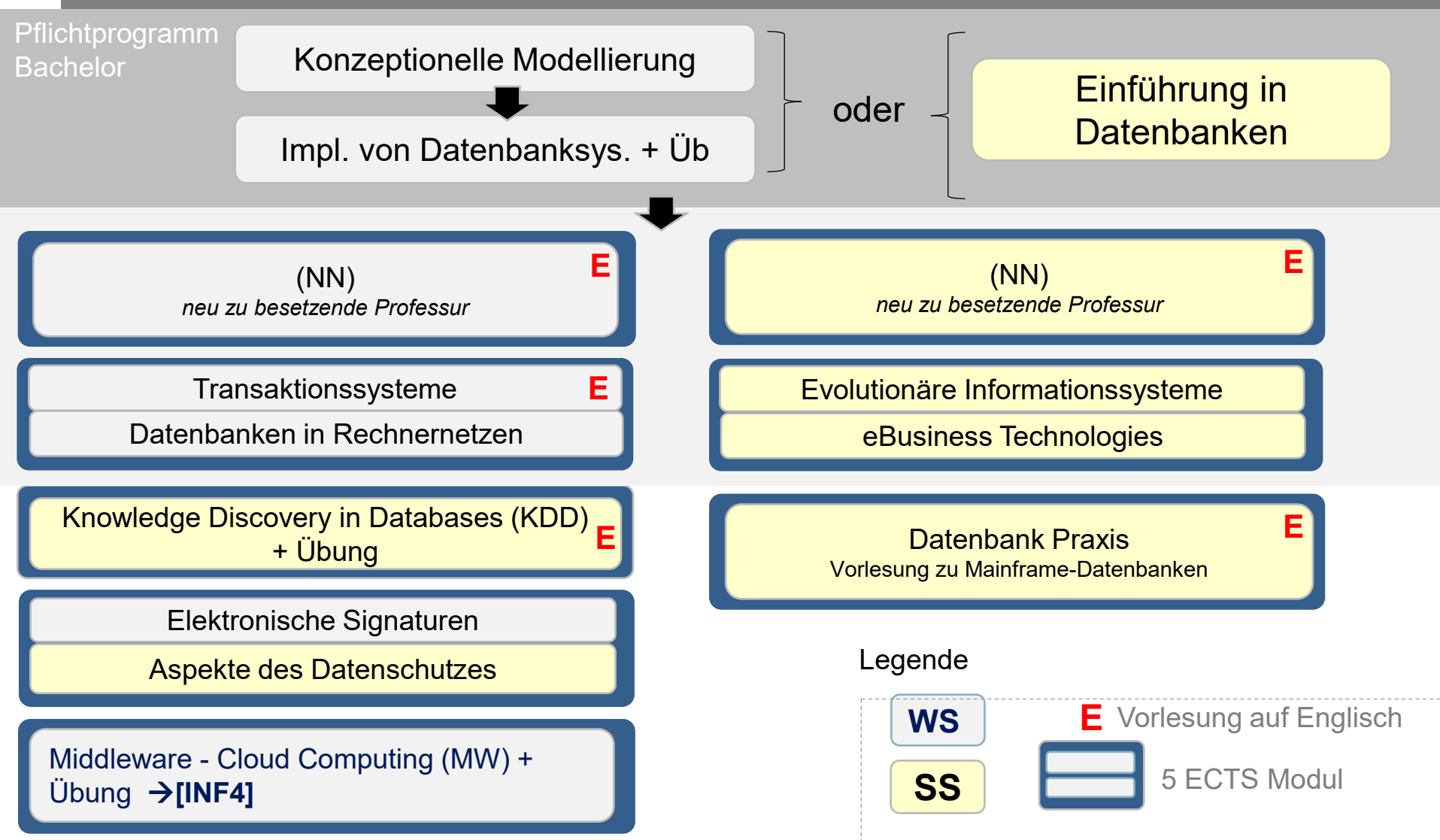
WS

SS

**E** Vorlesung auf Englisch



5 ECTS Modul





- **Hauptseminar:** „Big Data“ Seminar **WS** **SS**
  - Big Data – Technologien und Anwendungen
- SWAT („SWAT is a Web Application Tutorial“) (5 ECTS) **SS**
  - Übung / Praktikum als Vertiefung zur Vorlesung eBT
- Masterprojekt:  
Implementierungsprojekt Datenbanksysteme (10 ECTS) **SS**
  - Praktische Vertiefung der Vorlesung IDB
  - Im Bachelor verwendbar als: Implementierungspraktikum Datenbanksysteme
- Masterprojekt Datenmanagement (10 ECTS) **WS** **SS**
  - Praktikum mit individuell angepassten wechselnden Themen



<https://www.cs6.tf.fau.de/lehre/lehrveranstaltungen/>

# Jenseits der Module

Themengebiete und Forschungsschwerpunkte

Richard Lenz

Professur für Evolutionäres Datenmanagement /

Lehrstuhl für Informatik 6 Datenmanagement





- **Entwicklung effizienter und skalierbarer Datenbanksysteme**
  - Anfrageverarbeitung auf moderner Hardware
  - Datenmanagement in der Cloud
  - Transaktionsverarbeitung
  - Indexstrukturen
- **Evolutionäres Datenmanagement**
  - Datenintegration
  - Datenqualitätsmanagement
  - Schemaevolution
  - Data Lake Management

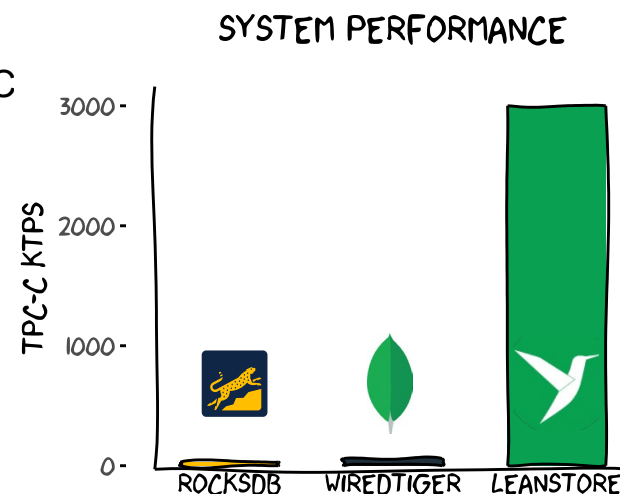
## Projekt leanstore.io

A High-Performance  
Storage Engine for Modern Hardware

(Kooperation mit Prof. Leis – TUM)

- high-performance OLTP storage engine optimiert für many-core C
- **Projekt Pufferverwaltung und Datenzugriff**  
Zugriffswege in Datenbanksystemen  
unter Einbeziehung moderner Hardware
  - Hauptspeicherverwaltung
  - Seitenersetzungsstrategien
  - Asynchroner Datenzugriff

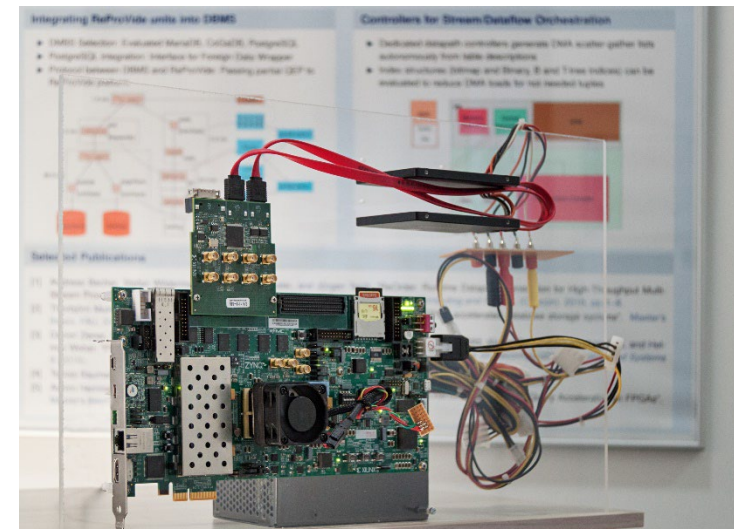
→ **Demian Vöhringer**



- **Projekt ReProVide:**

Query Optimisation and Near-Data Processing on Reconfigurable SoCs for Big Data Analysis

- Ausnutzung neuer Hardware für die Anfrageoptimierung in Datenbanksystemen
- Verarbeitung von Strömen und relationalen Anfragen
- Kooperation mit Lehrstuhl für Informatik 12 (Hardware-Software-Co-Design, Prof. Dr. **Jürgen Teich**)
- Abschlussarbeiten und Projekte offen



➔ **Maximilian Langohr**



---

## Speichereffizientes Datenmanagement

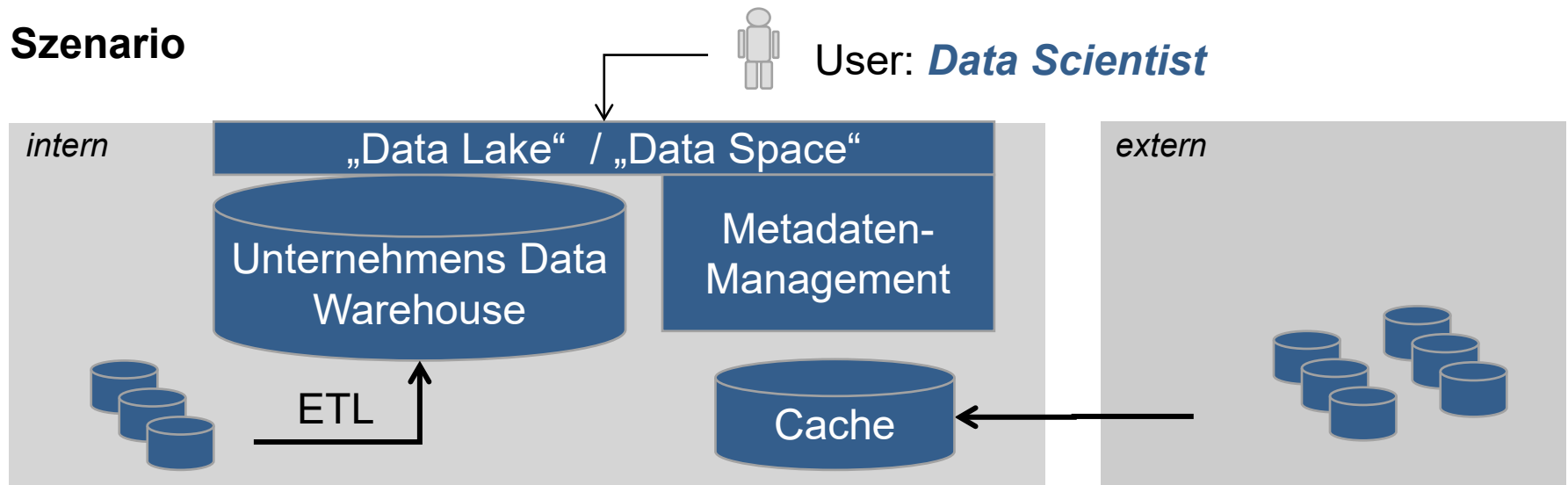
- **Projekt FST (Frequent Sub Trees):**
  - Projekt zur Optimierung von Stringkompressionen unter Ausnutzung von Baumstrukturen

→ **Dominik Probst**

Unternehmen unterliegen einem ständigen Wandel  
IT ist Teil der Unternehmen und sollte den Wandel nicht behindern

IT-Unterstützung für die  
**"Lernende Organisation"**

## Szenario





## Anfragegetriebenes offenes Datenmanagement für dynamisch wachsende Unternehmensdatenbestände

- Analyse von Anfrageprotokollen zur Verbesserung der Datenintegration
- Projekt: **OCEAN**  
Minimalinvasive Infrastruktur zur Query-Log-Analyse  
(→ Andreas Wahl )



- Nachfolgeprojekt: **PHAROS**  
Partielle Schemaextraktion  
mit Semantic-Web-Technologien  
→ **David Haller**





---

## Umgang mit unbekanntem Datenquellen Data Profiling und Schema-Inferenz

- **Projekt SIML:**
  - "Schema Inference and Machine Learning"
  - Interaktive Schema-Inferenz bei unbekanntem Sensordatenströmen
    - Erkennung von unbekanntem Sensordatenströmen aus Kraftwerken
    - Changepoint-Analyse

→ **Lucas Weber**



---

## Datenqualitätsmonitoring

- **Projekt DQ-Ontology:**

- Systematisches Management und Dokumentation von Datenqualitätsregeln und Maßnahmen in Reporting Data Lakes
- Kooperation mit SIEMENS HEALTHINEERS

→ **Jan-Lucas Deinhardt**





Vielen Dank ...



Fragen gerne an: [richard.lenz@fau.de](mailto:richard.lenz@fau.de)