



Vorlesung Implementierung von Datenbanksystemen

15. Epilog – Zusammenfassung und Ausblick

Prof. Dr. Klaus Meyer-Wegener
Wintersemester 2019/20

■ **Datenbanksystem**

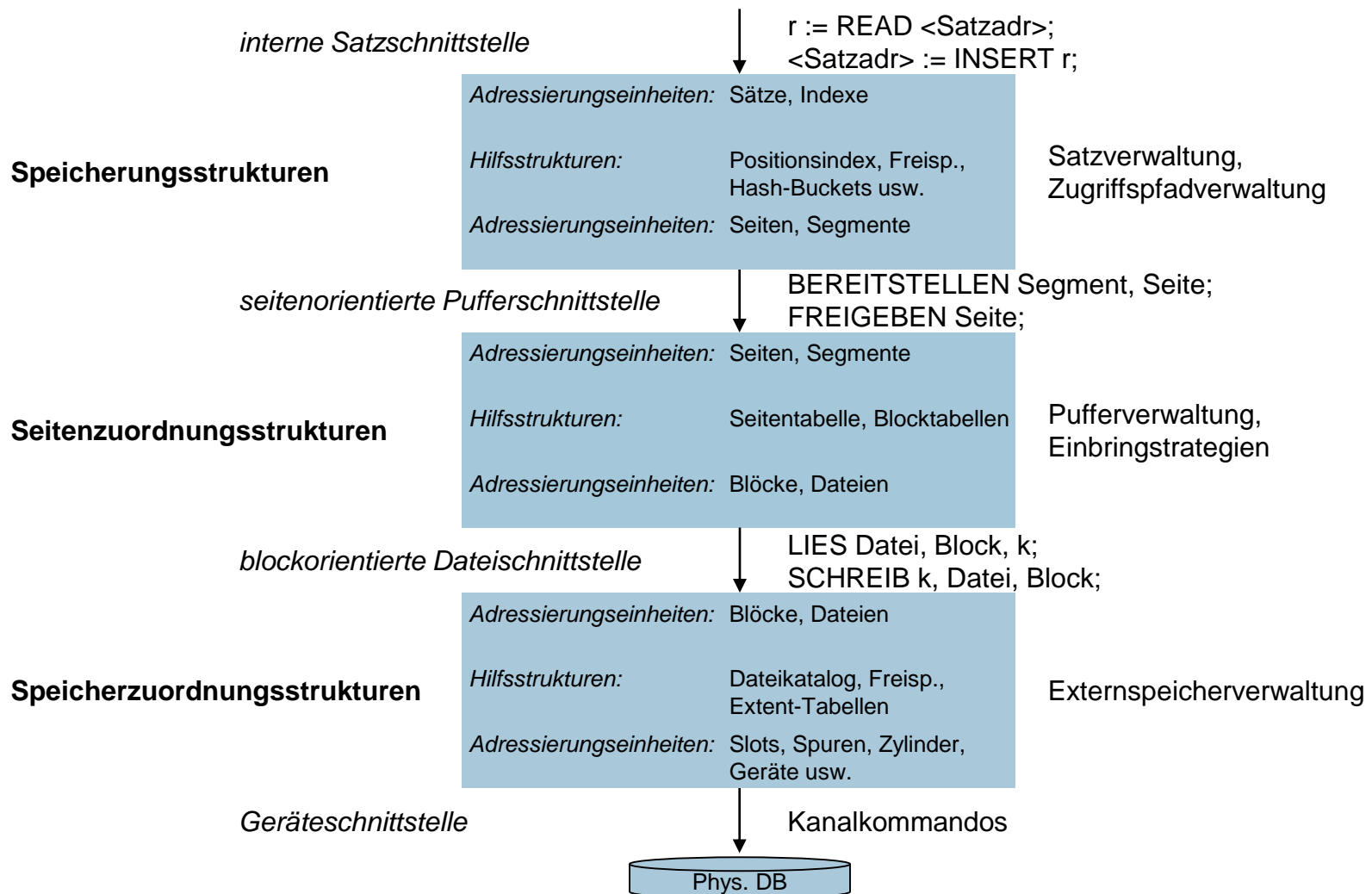
- Zentrale IT-System-Komponente in fast jedem Unternehmen
- Systemtechnische Grundlage des eCommerce / eBusiness / e???

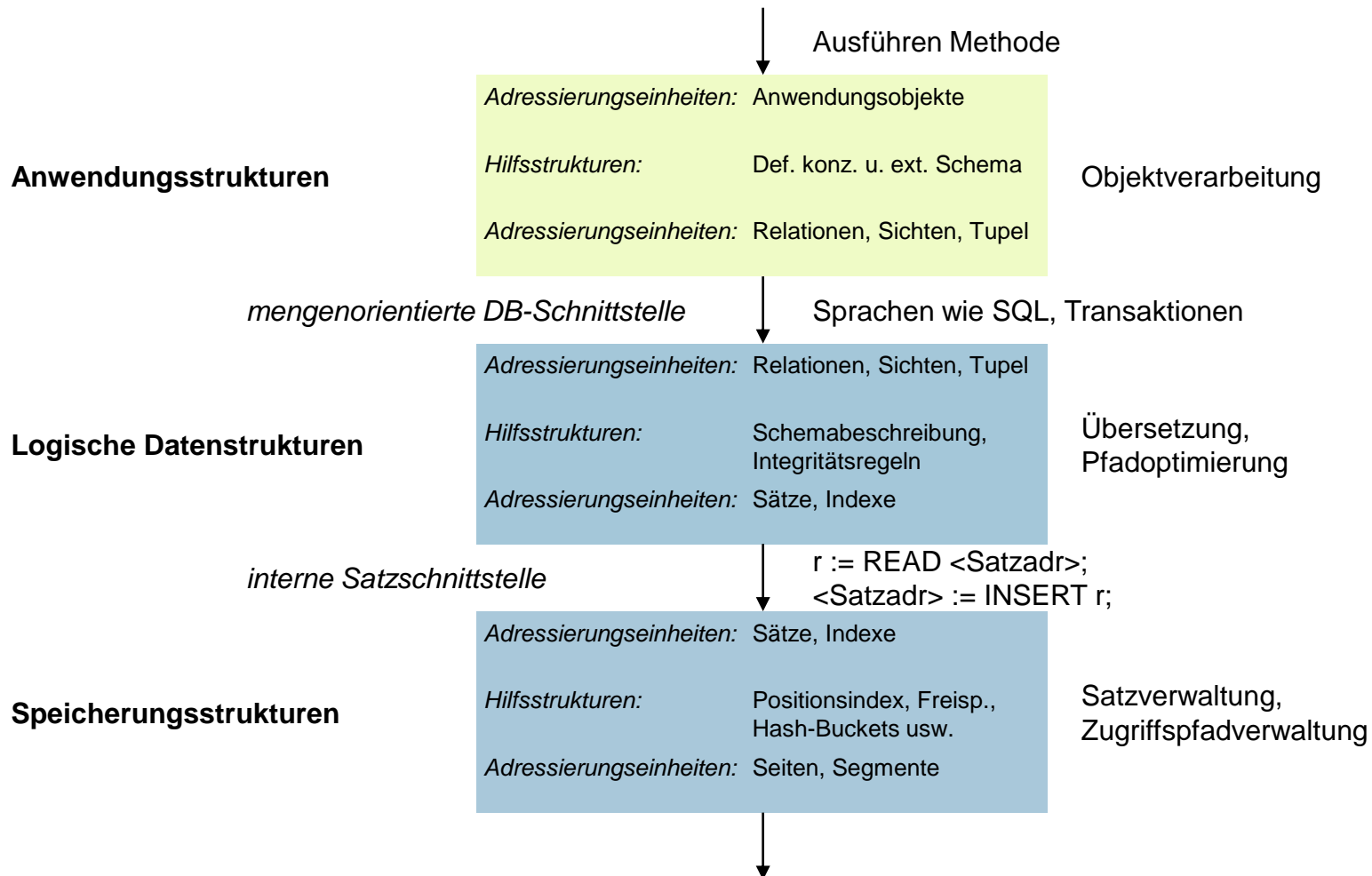
■ **Architektur von Datenbanksystemen**

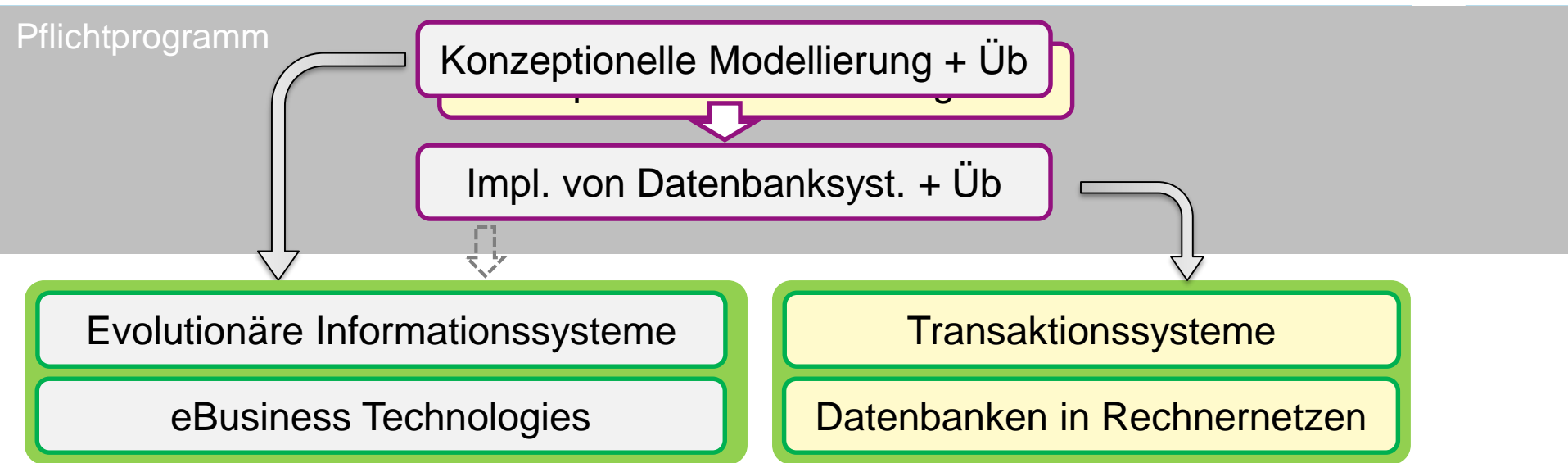
- Strukturierung (Schichtenmodell, Komponentensichtweise)
- Schwerpunkt auf generischer Softwareentwicklung (Anwendungsneutralität)
- Definition / Erfüllung von Schnittstellen
- Wechselwirkungen beteiligter Komponenten
- Speicher-, Zugriffs- und Verarbeitungsschicht

■ **Transaktionsverarbeitung**

- Synchronisation gleichzeitiger Zugriffe auf gemeinsame Daten
- Protokollierung und Wiederherstellung





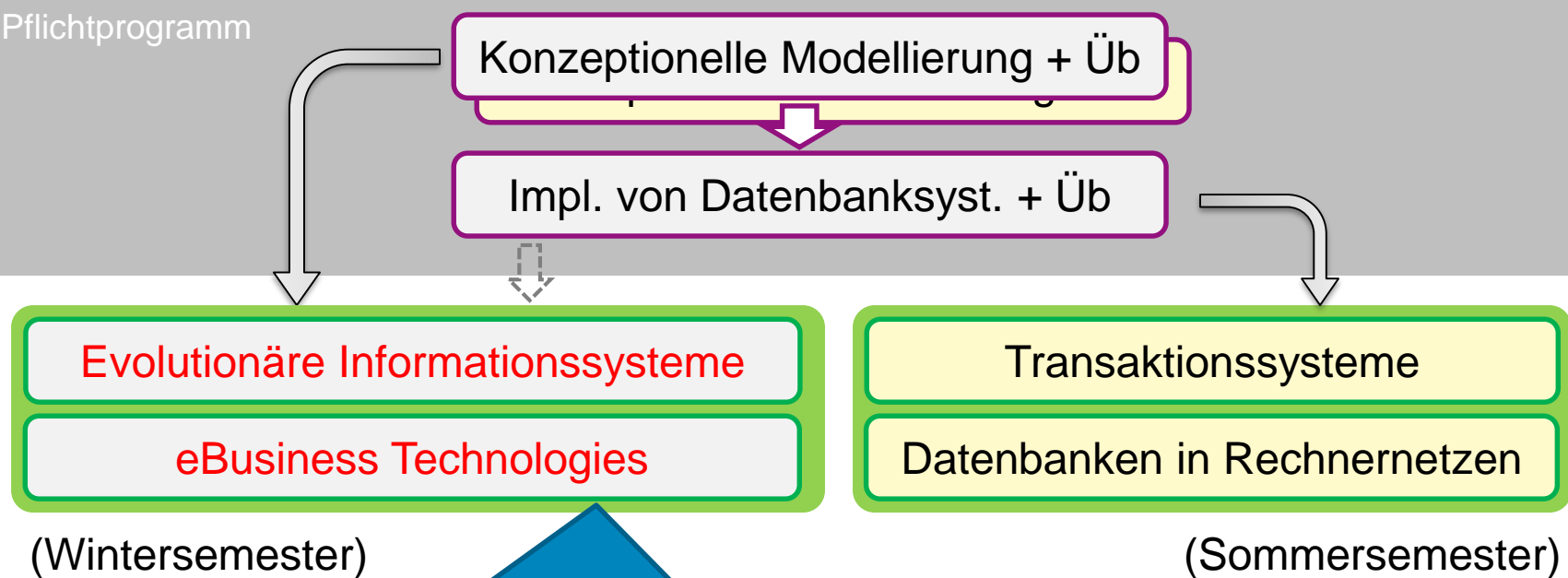


Module im Bachelor oder Master wählbar

- jeweils 5 ECTS

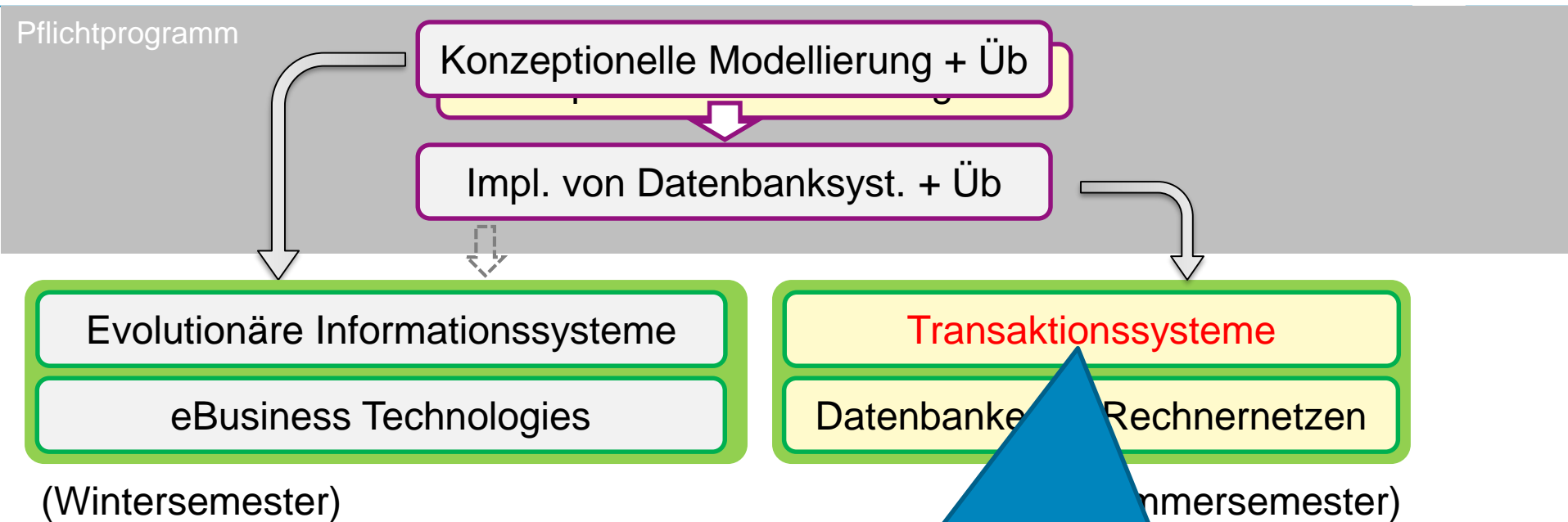
SS

WS



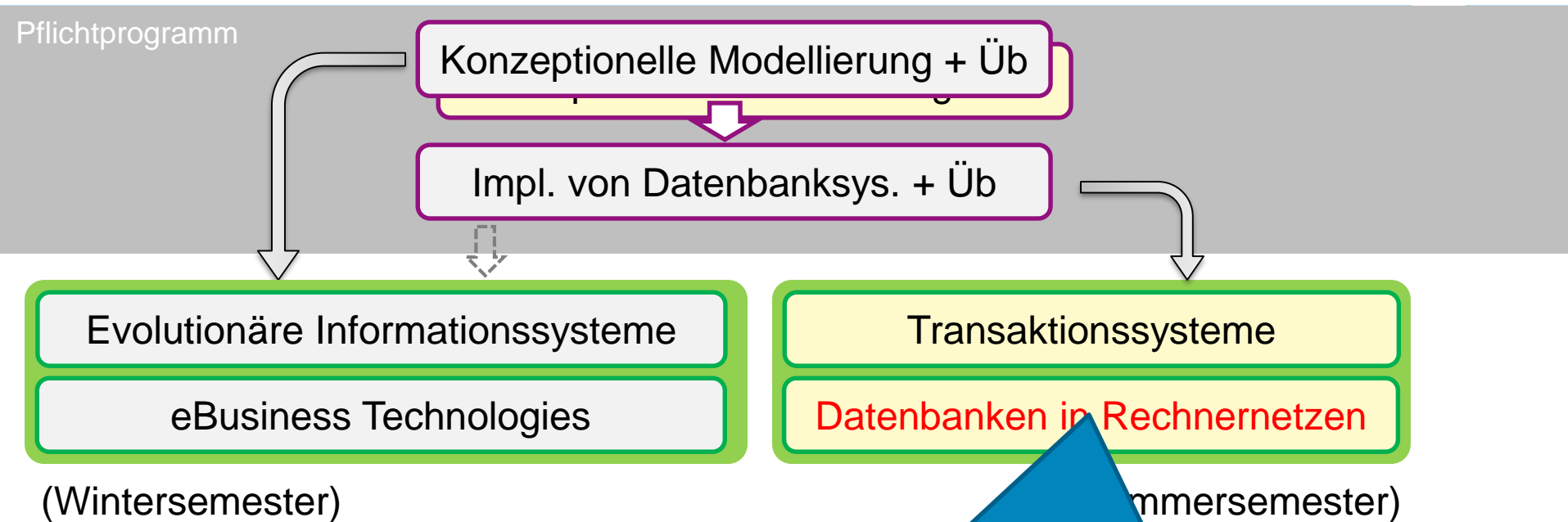
EIS und eBT:

- Einsatz von Datenbanksystemen in Unternehmen
- Datenbank-zentrierte Web-Anwendungen
- Transparenter Zugriff auf Datenbanken
- Umgang mit sich änderndem Bedarf



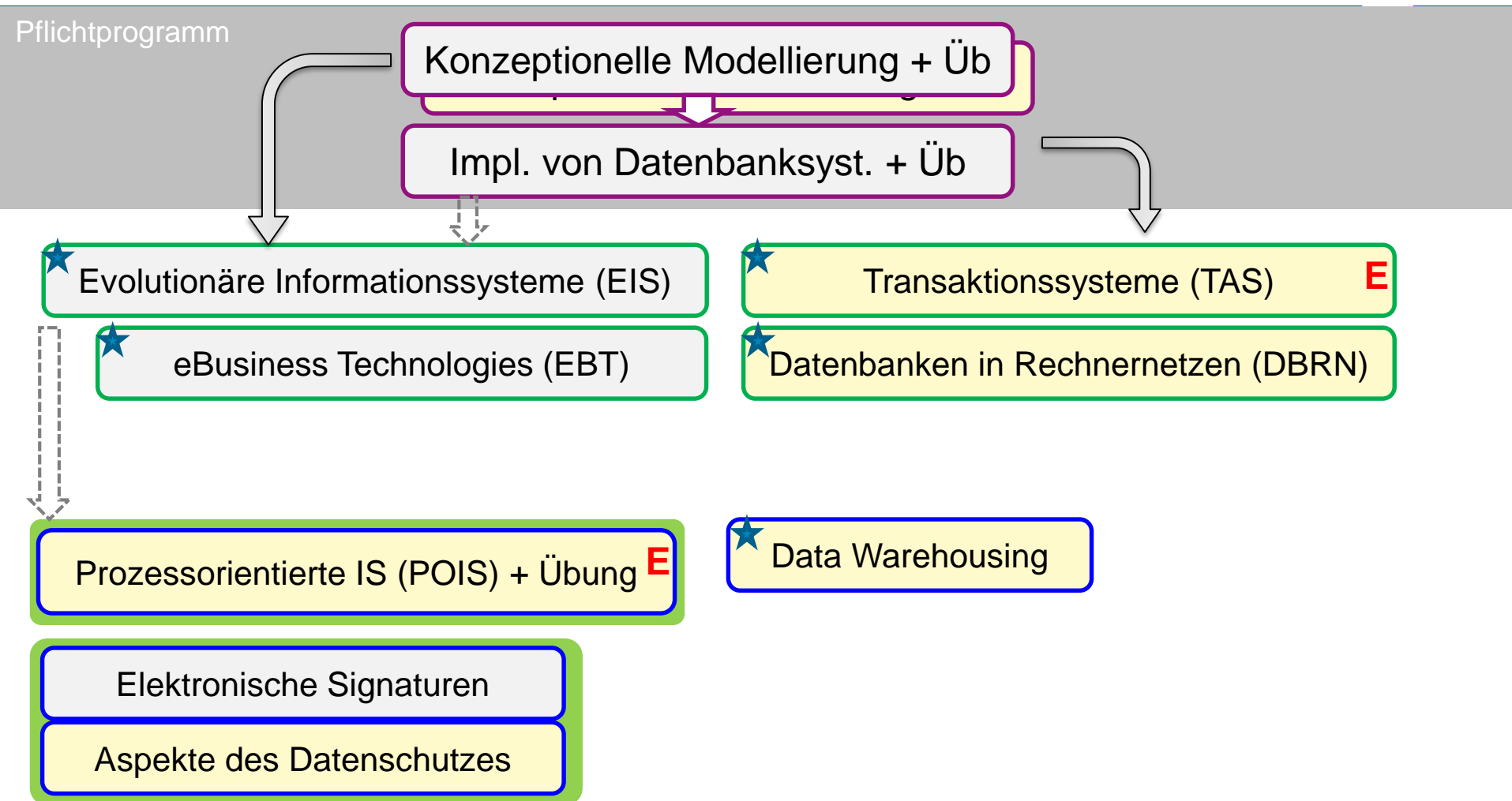
Transaktionssysteme:

- Fehlerbehandlung in Datenbanksystemen
- ACID-Transaktionen
- Erweiterte Transaktionsmodelle
- Synchronisation und Wiederherstellung im Detail



Datenbanken in Rechnernetzen:

- Verteilte Datenbanken
- Verteilte Anfrageverarbeitung
- Datenreplikation
- Heterogene Datenbanken → Föderative Datenbanken
- Parallele Anfrageverarbeitung: Umgang mit "Big Data"



★ Je 2 Veranstaltungen im selben Semester kombinierbar zu einem 5-ECTS-Modul

E = Englisch

- **Bachelor-Seminar**
 - Z.B. im Wintersemester 2019/20: Big Data
- **Praktikum**
 - Z.B. im Sommersemester 2020: SWAT ("SWAT is a Web Application Tutorial" ☺ ; siehe nächste Folie)
- **Bachelorarbeit**
 - Angebote stets auch im UnivIS
- **Master-Seminar**
 - Z.B. im Sommersemester 2020: Big Data und Topologische Datenanalyse
- **Projekt**
 - Z.B. im Sommersemester 2020: individuell und projektbezogen (UnivIS)
- **Masterarbeit**
 - Angebote stets auch im UnivIS

■ Task:

- Entwurf und Implementierung einer typischen Web-Applikation
- Kreatives Arbeiten im Team
- Agile Softwareentwicklung
- Verwendung von aktuellen Technologien
- Moderne Programmiertechniken

■ Tools & Techniques:

- Apache Jena; Spring; Vaadin
- Java; SPARQL; RDF; SCRUM

vaadin }>

 **Jena**

 **spring**[®]

Topic SS2020: Graphbasierte Datenspeicherung und -auswertung

Datenbanksysteme

Effiziente Speicherung und Verarbeitung von Daten

Evolutionäre Informationssysteme

Umgang mit Veränderung

Architektur evolutionärer Informationssysteme

Prozessunterstützung

Datenqualität

Methoden und Werkzeuge zur
Verbesserung der Datenqualität
in Unternehmen

Datenstromsysteme und Ereignisverarbeitung

Umgang mit zeitbehafteten Datensätzen

BIG-DATA-ASPEKTE

Datenbanksysteme

Effiziente Verarbeitung von Daten

Neue Strategien für neue Anforderungen

Evolutionäre Informationssysteme

Umgang mit Veränderung
Architektur evolutionärer Informationssysteme
Prozessunterstützung

Immer schnellere Reaktion auf neue Erkenntnisse

Datenqualität

Methoden und Werkzeuge zur Verbesserung der Datenqualität in Unternehmen

Traditionelle Methoden der Qualitätssicherung versagen bei externen Datenquellen.

Datenstromsysteme und Ereignisverarbeitung

Umgang mit zeitbehafteten Datensätzen

Aktuelle zeitbehaftete Daten müssen mit Unternehmensdaten verknüpft werden.

- **ReProVide:**

Query Optimisation and Near-Data Processing on Reconfigurable SoCs for Big Data Analysis (DFG-Projekt)

➔ **Lekshmi B.G., M.Sc., M.Phil.**

- Ausnutzung neuer Hardware für die Anfrageoptimierung in Datenbanksystemen
- Kooperation mit Lehrstuhl für Informatik 12 (Hardware-Software-Co-Design, Prof. Teich)

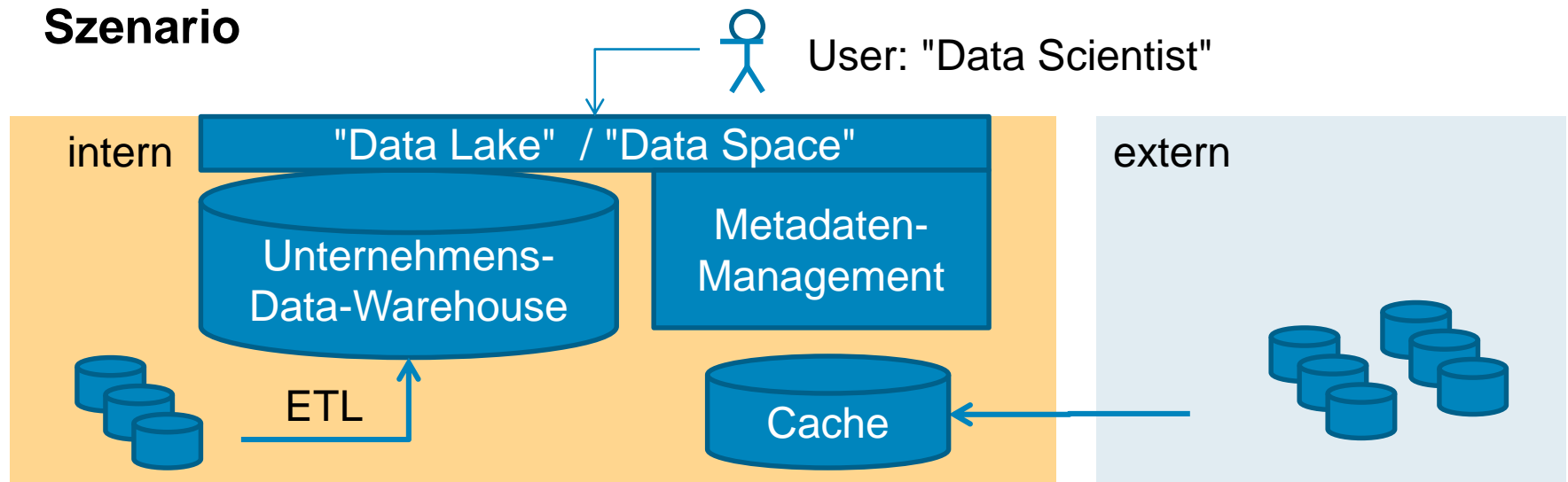


Von Altera Corporation - Altera Corporation, CC BY 3.0,
<https://commons.wikimedia.org/w/index.php?curid=6642224>

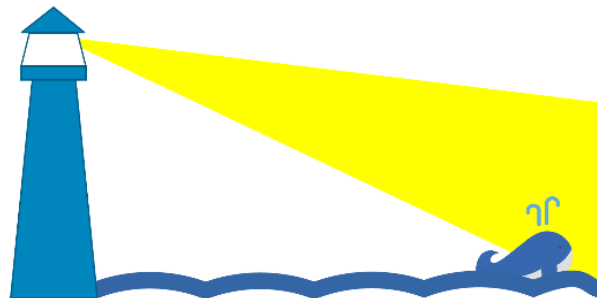
Unternehmen unterliegen einem ständigen Wandel
IT ist Teil der Unternehmen und sollte den Wandel nicht behindern

IT-Unterstützung für die
"Lernende Organisation"

■ Szenario



- **Anfrage-getriebenes offenes Datenmanagement für dynamisch wachsende Unternehmensdatenbestände**
 - Anfrage-getriebene bedarfsorientierte Schemaentwicklung
→ **Peter Schwab**
 - **PHAROS:** Learning schemata from SQL Queries and describing them in a knowledge graph
→ **David Haller**



■ Warum Datenqualität ?

Datenqualitätsprobleme verursachen in USA jährlich Kosten in Höhe von \$ 600 000 000 000.

Quelle: *Wayne Eckerson:*

Data Quality and the Bottom Line: Achieving Business Success through a Commitment to High Quality Data.

The Data Warehouse Institute. Repost Series. 2002

Das "Jahr-2000"-Problem war ein Datenqualitätsproblem, das geschätzte Kosten von \$ 1 500 000 000 000 verursacht hat.

Quelle: *L.P. English:*

Improving Data Warehouse and Business Information Quality.
Wiley & Sons. 1999

- **Projekte**

- **SIML – Schema Inference and Machine Learning**

Knowledge Discovery; Functional Dependency; topologische Datenanalyse; Clustering; unsupervised Learning; Neural-Network Theory

➔ **Luciano Melodia**



Datenstromsysteme:

- Ich will das verarbeiten, was sowieso kommt.

Datenbanksysteme:

- Es kommt genau so viel, wie ich möchte – aber nicht von allein.



- **Data-Stream Management System (DSMS)**

- Keine Speicherung von Daten, Auswertung "on the fly", geringe Latenz
- **Sensornetzwerke** als Datenquelle
- Identifikation von Ereignissen (Complex-Event Processing, CEP)
- Data-Stream Application Manager (DSAM)

- **Projekte:**

- EFRE-**E|ASY-Opt** Teilprojekt
Digitalisierung: **Industrie 4.0 und Predictive Maintenance**
➔ Vorausschauende Wartung der anfallenden Sensordaten bei
Fertigungsanlagen mittels Data Mining bzw. Machine Learning
➔ **Melanie Sigl**

- **Informatik in den Geistes- und Sozialwissenschaften**
 - Analyse von Texten und Bildern, umfassende Annotationen
 - Wissensbanken
 - Multimedia

- **Projekt:**
 - **WissKI:**
Wissenschaftliche Kommunikationsinfrastruktur;
in der AG Digital Humanities (Prof. Dr. Günther Görz);
Content Management (Drupal), RDF, OWL, Multimedia
→ **Corina Lehmann, Peggy Große** (AG Digital Humanities, GNM)

- **Vorauss. am Freitag, dem 14. Februar 2020**
 - Uhrzeit und Räume stehen noch nicht fest.
- **Dauer: 90 Minuten**
- **Viel Multiple Choice**
- **Keine Hilfsmittel!**
- **Fragestunde heute, hier**
- **Fragen in der Vorbereitung werden auch im StudOn-Forum beantwortet.**

Viel Erfolg und alles Gute!!!