



Technische Hochschule
Ingolstadt

Fakultät für Elektrotechnik
und Informatik

*Zukunft in
Bewegung*

IT-Integrations- und Migrationstechnologien

Heterogene Systeme

Prof. Dr. Bernd Hafenrichter 01.10.2023

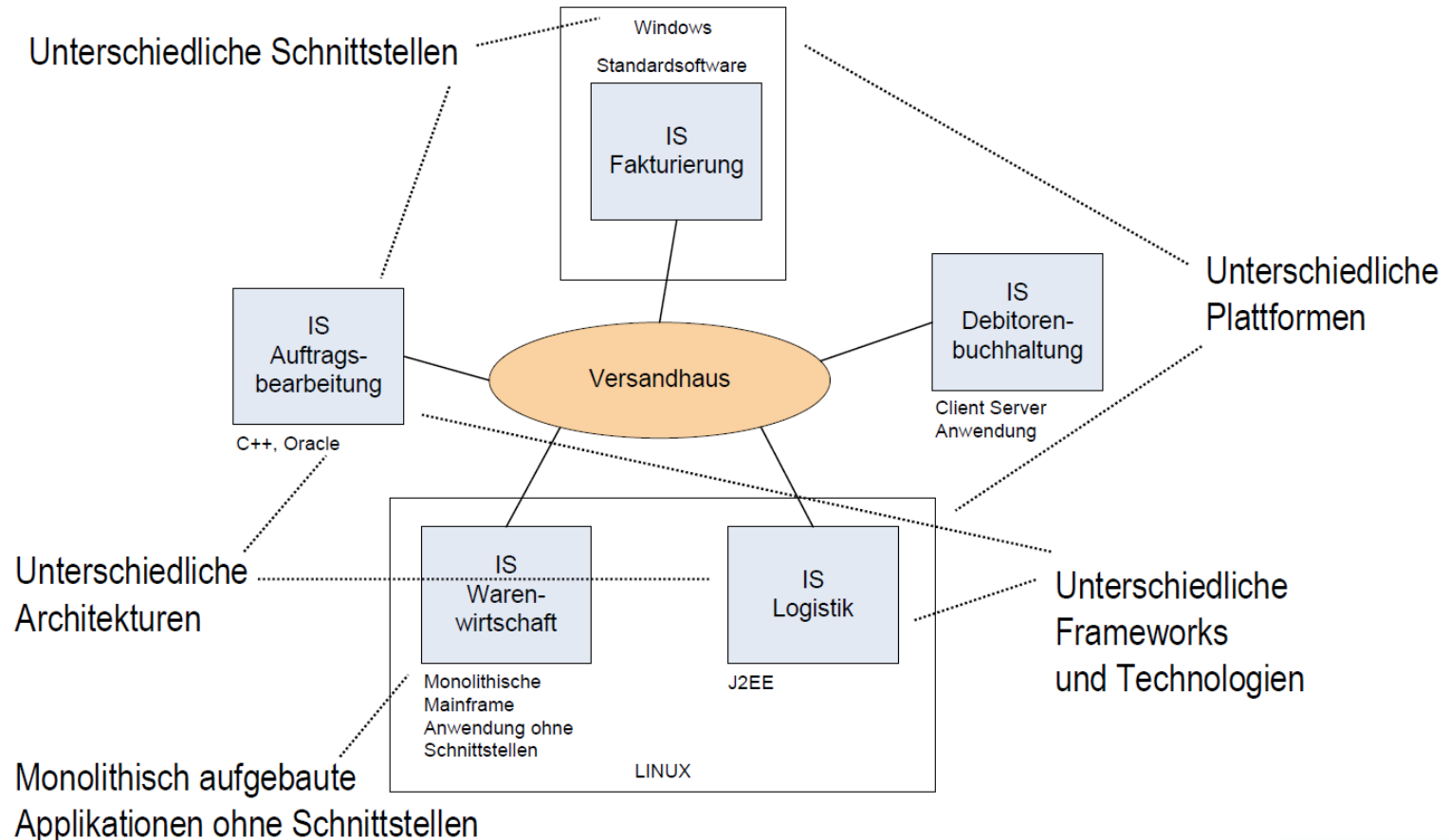




Motivation

- IT-System unterscheiden sich sehr stark in der eingesetzten Technologie, Betriebsumgebung, Datenbank, Protokolle usw.
- Die Integration heterogener System gestaltet sich als sehr schwierig, da erst eine gemeinsame (technologische) Kommunikationsbasis geschaffen werden muss
- Eine Integrationslösung muss Mittel- und Werkzeuge bereitstellen um diese Aufgabe gerecht zu werden
- Evtl. sind Anpassungen in den beteiligten Systemen notwendig

Motivation





Föderierte Systeme

- Vgl. „BRD“ oder „EU“ (Teilstaaten, Gesamtzweck)
- Föderation und deren Charakterisierung über
 - Grad der Autonomie
 - Heterogenität der beteiligten Entitäten (Systeme, Länder, ...)
- Analog: Föderation von IT-Systemen, klassifiziert nach
 - Grad der **Autonomie** der integrierten Komponenten
 - Grad der **Heterogenität** zwischen den Komponenten
 - Grad ihrer **Verteilung**



Föderierte Systeme

- 1 Autonomie
 - **Entwurfsautonomie:**
Unabhängiger Entwurf des Aufbaus (eigene Architektur, bspw. des Datenmodells)
 - **Kommunikationsautonomie:**
Eigene Entscheidung, mit wem kommuniziert werden soll
 - **Ausführungsautonomie:**
Eigene Entscheidung über Ausführung und „Scheduling“ eingehender Anfragen

vgl. [CONRAD et al., 2007, S. 11 ff]

Föderierte Systeme

2 Heterogenität

- **Syntaktische Heterogenität**

- Technische Heterogenität
- Schnittstellen-Heterogenität

- **Datenmodellbasierte Heterogenität**

Unterschiedliche Möglichkeiten in der Abbildung: z. B. OO vs. ERM

- **Logische Heterogenität**

- **Semantische Heterogenität**

Bedeutung von Entitäten, z. B. „Kunde“ in IT-System A und B das Selbe?

- **Schemabasierte Heterogenität**

Modellierung von Entitäten mit verschiedenen Elementen eines Modells

- **Strukturelle Heterogenität**

Selbe Bedeutung, selbes Modell, semantisch homogen aber unterschiedlich strukturiert (d. h. innerhalb der einzelnen Datensätze)

vgl. [CONRAD et al., 2007, S. 11 ff]



Föderierte Systeme

3 Verteilung

- **Verteilung (physisch) von**
 - Daten(quellen)
 - IT-Systemen
- **Weitere Aspekte hierbei:**
 - Physische (globale) Verteilung über Netzwerke/Internet
 - Verschwimmende Grenzen der Verteilung durch Virtualisierung (vgl. z. B. „Cloud-Computing“, „Grid-Computing“)



Syntaktische Heterogenität

Technische Heterogenität

Unterschiede auf der Ebene der technischen Infrastrukturen wie DBMS, Hardwareplattformen, Betriebssysteme und Netzwerkkomponenten

Beispiel:

Lagerverwaltungssystem (LVS) installiert auf MVS-Großrechner
E-Commerce-System (ECS) J2EE-Application auf Linux

Problem: Zeichensatz

LVS verwendet EBCDIC

ECS verwendet Unicode

Für eine Integration ist eine Zeichenkonvertierung notwendig

Syntaktische Heterogenität

Technische Heterogenität

- Unterschiedliche Darstellung desselben Sachverhalts
- Dezimalpunkt oder –komma
- Euro oder €
- Comma-separated oder tab-separated
- ASCII, Unicode oder EBCDIC
- Little Endian or Big Endian kodierung
- Notenskala 1-6 oder „sehr gut“, „gut“, ...
- Binärcodierung oder Zeichen
- Datumsformate (12. September 2006, 12.9.2006, 9/12/2006, ...)

Datenkonflikt durch unterschiedliche Verwendung von Datentypen.

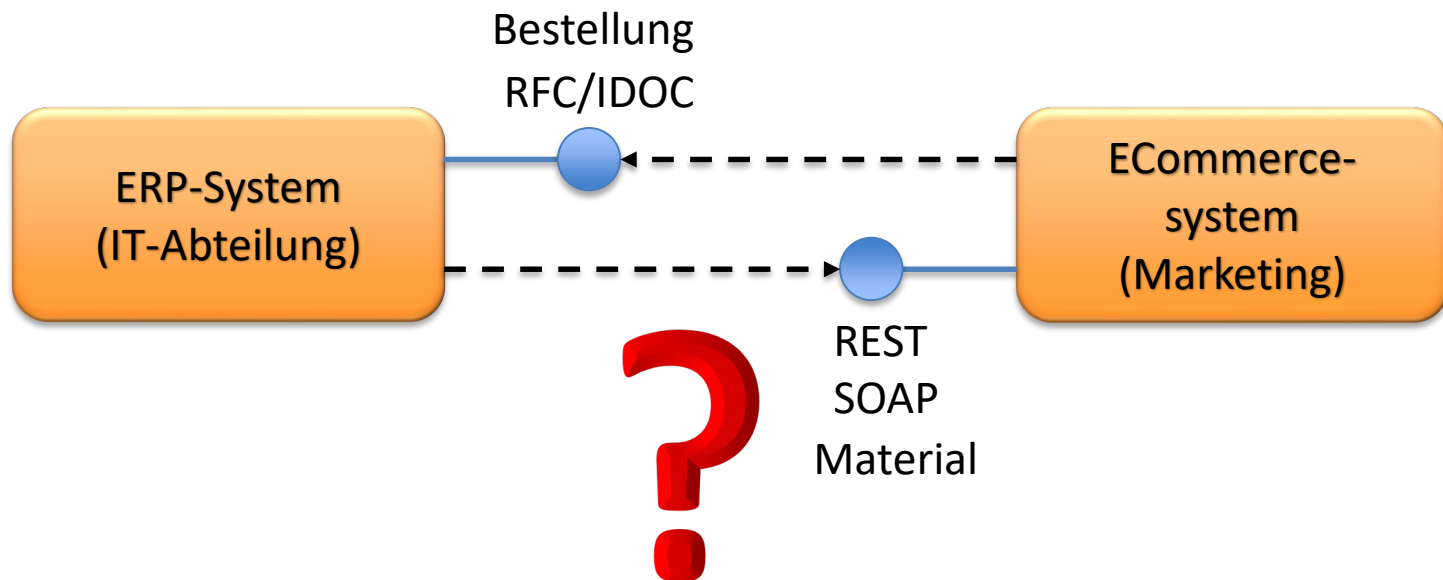
Tabelle: Mitarbeiter			
Nr	Name	Vorname	Datum
1	Mayer	Fritz	01.01.2014
2	Huber	Verena	01.10.2013

Tabelle: Mitarbeiter			
EMP_ID	Name	Vorname	Datum
1	Mayer	Fritz	20140101
2	Huber	Verena	20131001

Ebenen der Heterogenität – Syntaktische Heterogenität

Schnittstellenheterogenität:

Vorhandene technische Schnittstellen können aufgrund unterschiedlicher Protokoll nicht genutzt werden.



Datenmodellbasierte Heterogenität

Datenmodellbasierte Heterogenität

Unterschiede in der Datenmodellierung, ausgelöst durch Anwendung unterschiedlicher Modellierungskonzepte oder –Sprachen (z.B. relationales vs. objektorientiertes Datenmodell).

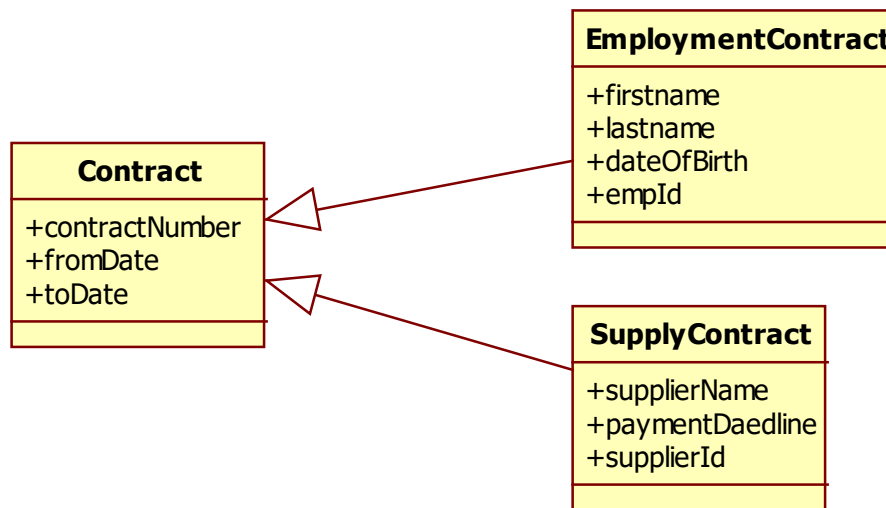
Mögliche Datenmodelle

- Hierarchisch (z.B. XML)
- Relational
- Domänen spezifische Modelle
- Objektorientierung

Datenkonflikte durch Auswahl unterschiedlicher Objekteigenschaften oder Einsatz uneinheitlicher Bezeichner, Datentypen, Wertebereiche und Skalierungen für semantisch vergleichbare Objekte.

Datenmodellbasierte Heterogenität

Anwendungsdomäne: Vertragsverwaltung (objektorientiert)



Datenmodellbasierte Heterogenität

Anwendungsdomäne: Vertragsverwaltung (relational)

SingleTable

Contract

contractType	contractNumber	fromData	toDate	firstname	lastname	dateOfBirth	empld	supplierName	paymentDeadline	supplierId

OneTablePerClass

Contract

contractNumber	fromData	toDate

Employee

contractNumber	fromData	toDate	firstname	lastname	dateOfBirth	empld

Supplier

contractNumber	fromData	toDate	supplierName	paymentDeadline	supplierId

Datenmodellbasierte Heterogenität

Anwendungsdomäne: Vertragsverwaltung (relational)

JoinedTable

Contract

contractNumber	fromData	toDate

Employee

contractNumber	firstname	lastname	dateOfBirth	empld

Supplier

contractNumber	supplierName	paymentDeadline	supplierId

Ebenen der Heterogenität

Logische Heterogenität

Semantische Heterogenität

Bedeutung von Entitäten, z. B. „Kunde“ in IT-System A und B das Selbe?

Schemabasierte Heterogenität

Modellierung von Entitäten mit verschiedenen Elementen eines Modells

Strukturelle Heterogenität

Selbe Bedeutung, selbes Modell, semantisch homogen aber unterschiedlich strukturiert (d. h. innerhalb der einzelnen Datensätze)

Logische Heterogenität

Semantische Heterogenität

Für ähnliche Konzepte existieren in verschiedenen Systemen unterschiedliche Bedeutungen, welche sich aus der zugrunde liegenden Geschäftslogik ableiten und in der Programmlogik verankert sind.

Tabelle: Mitarbeiter				
Nr	Name	Vorname	Datum	Geschlecht
1	Mayer	Fritz	01.01.2014	0
2	Huber	Verena	01.10.2013	1
3	Hübner	Josef	01.10.2013	2

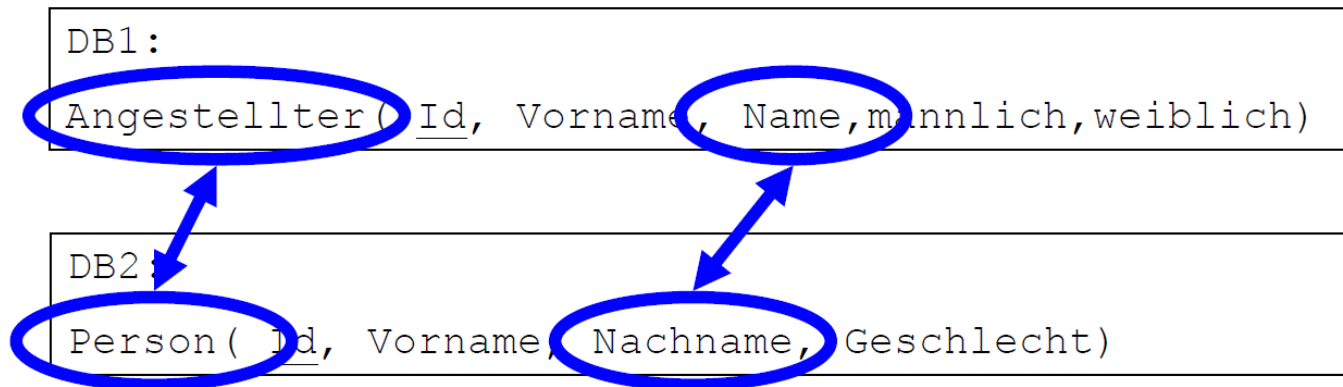
Tabelle: Employees				
EMP_ID	EMP_LastName	EMP_FirstName	EMP_DateOfEntry	EMP_SEX
1	Mayer	Fritz	20140101	M
2	Huber	Verena	20131001	F
3	Hübner	Josef	20131001	????

Logische Heterogenität

Semantische Heterogenität

Für ähnliche Konzepte existieren in verschiedenen Systemen unterschiedliche Bedeutungen, welche sich aus der zugrunde liegenden Geschäftslogik ableiten und in der Programmlogik verankert sind.

Synonyme (= Verschiedene Worte gleiche Bedeutung)

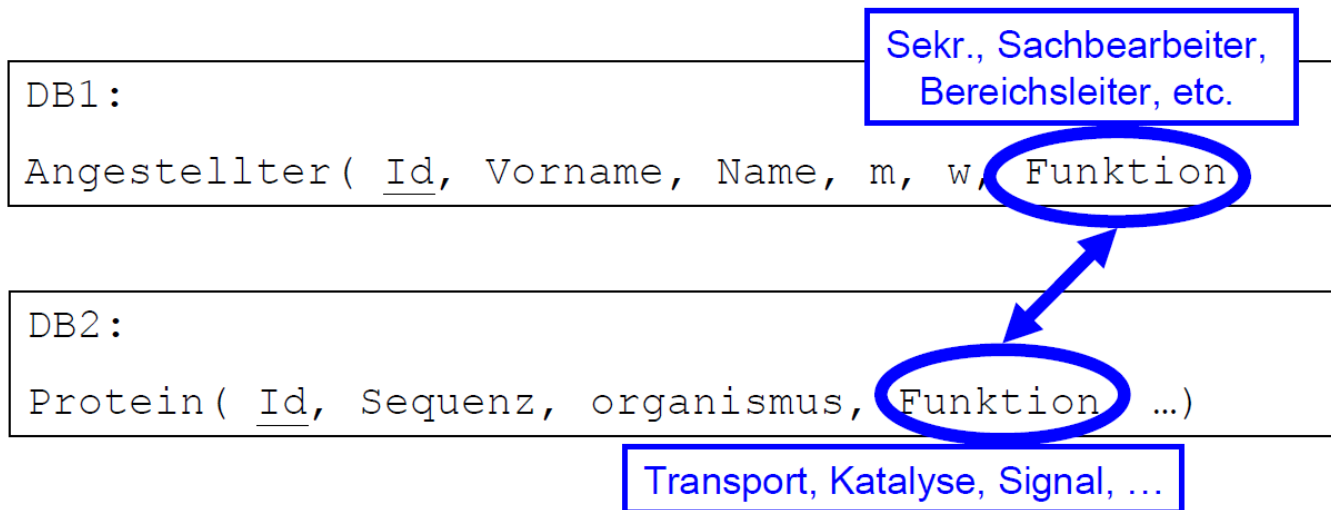


Logische Heterogenität

Semantische Heterogenität

Für ähnliche Konzepte existieren in verschiedenen Systemen unterschiedliche Bedeutungen, welche sich aus der zugrunde liegenden Geschäftslogik ableiten und in der Programmlogik verankert sind.

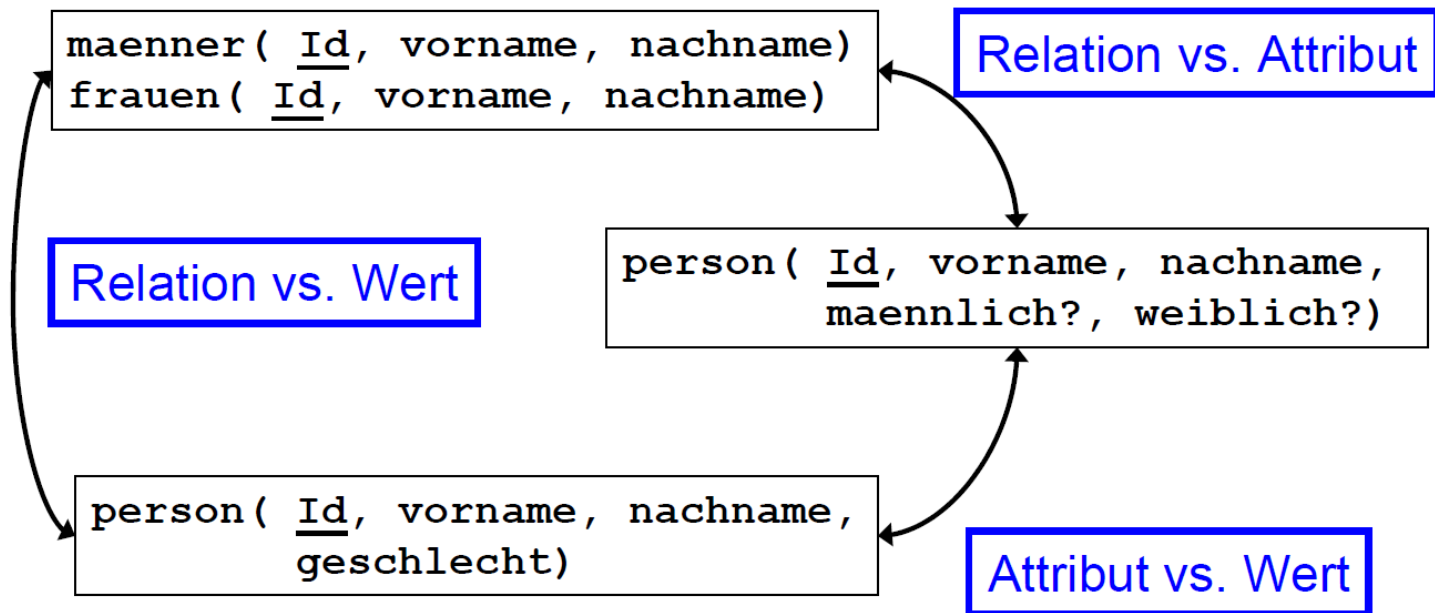
Homonyme (= Gleiche Worte verschiedener Bedeutung)



Logische Heterogenität

Schemabasierte Heterogenität

Reale Objekte werden mit unterschiedlichen Elementen einer Modellierungssprache dargestellt



Ulf Leser, Wissensmanagement in der Bioinformatik, Humboldt Universität



Logische Heterogenität

Strukturelle Heterogenität

Definition 1: Elemente haben die gleiche Bedeutung, sind semantisch Homogen, aber unterschiedliche strukturiert.

Beispiel: wird in einem Schema der Geburtsort mit PLZ bei der Person erfasst, in einem anderen Schema durch eine separate Tabelle mit 1:1-Beziehung zur Person).

Ebenen der Heterogenität – Syntaktische Heterogenität

Beispiel

Datenkonflikt durch uneinheitliche Verwendung der Spaltenname.

Tabelle: Mitarbeiter			
Nr	Name	Vorname	Datum
1	Mayer	Fritz	01.01.2014
2	Huber	Verena	01.10.2013

Tabelle: Employees			
EMP_ID	EMP_LastName	EMP_FirstName	EMP_DateOfEntry
1	Mayer	Fritz	01.01.2014
2	Huber	Verena	01.10.2013

Datenkonflikt durch unterschiedliche Verwendung von Datentypen.

Tabelle: Mitarbeiter			
Nr	Name	Vorname	Datum
1	Mayer	Fritz	01.01.2014
2	Huber	Verena	01.10.2013

Tabelle: Employees			
EMP_ID	EMP_LastName	EMP_FirstName	EMP_DateOfEntry
1	Mayer	Fritz	20140101
2	Huber	Verena	20131001

Wie kann die Heterogenität durch Integration überwunden werden?

Beispiel

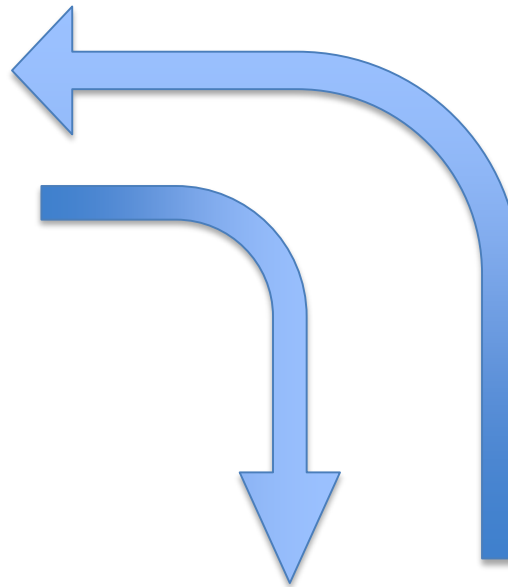
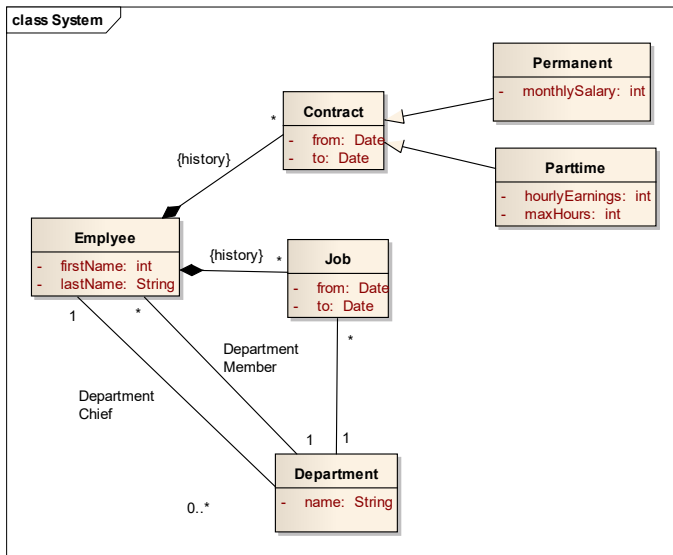


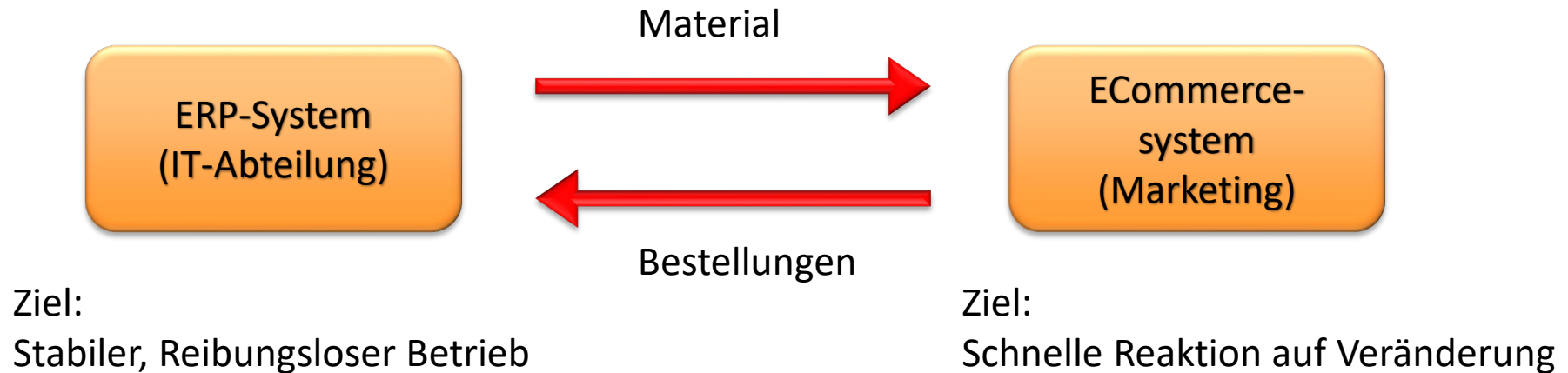
Tabelle: Employees					
EMP_ID	EMP_LastName	EMP_FirstName	EMP_DateOfEntry	EMP_Salary	EMP_DepartmentName
1	Mayer	Fritz	20140101		
2	Huber	Verena	20131001		

Ebenen der Heterogenität

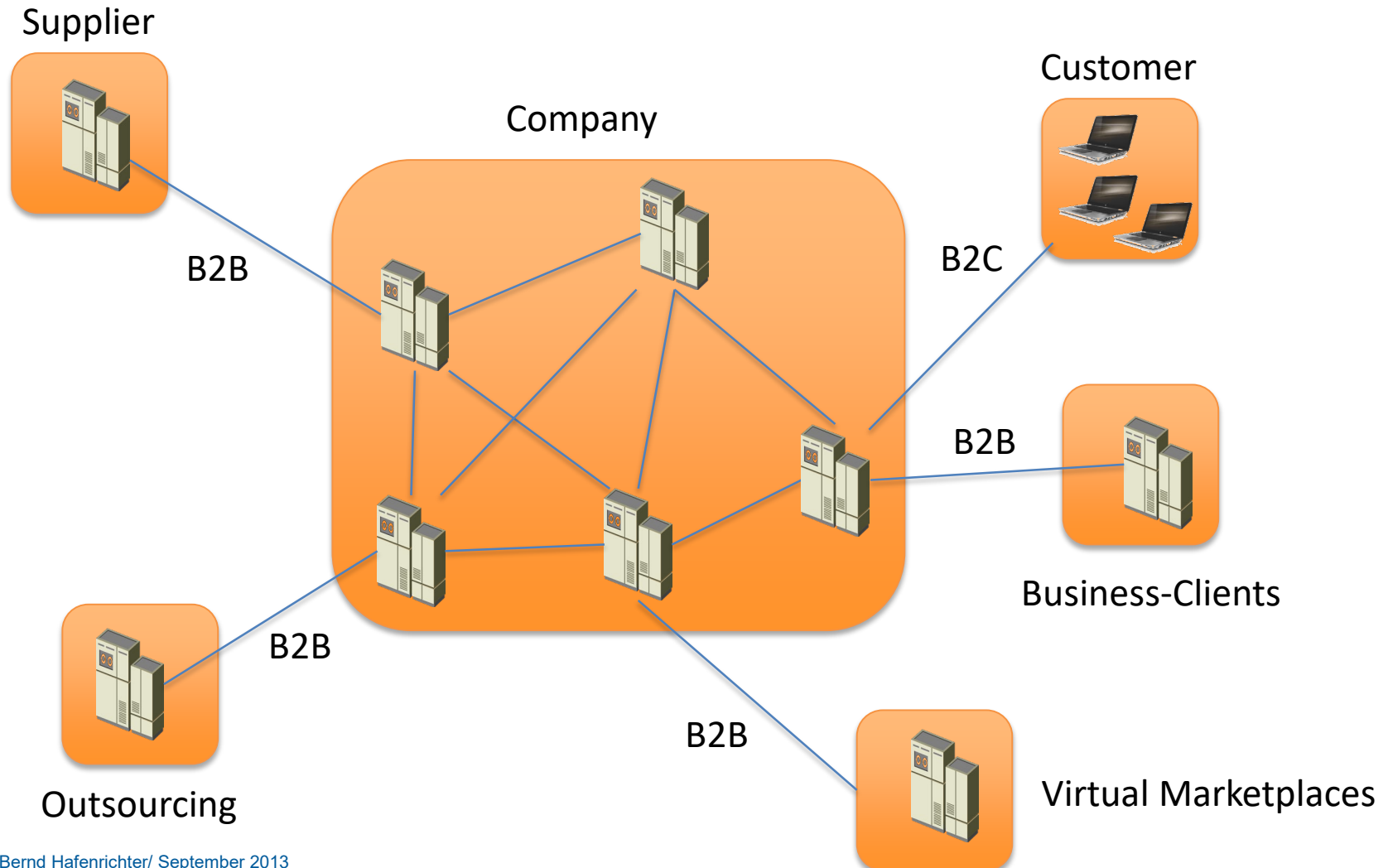
Autonomie und Verteilung

Systeme sind räumlich getrennt und unterstehen unterschiedlichen Verantwortlichkeiten.

Beispiel



Heterogenität im Verantwortungsbereich





Heterogenität im Verantwortungsbereich

- **Besondere Herausforderungen an die Integration**
 - **Sicherheit:**
 - Verschlüsselung, Integrität, Authentizität
 - End-2-End Verschlüsselung
 - **Kommunikationsprotokolle:**
 - Protokolle müssen über Firewall und Firmengrenzen hinaus funktionieren
 - **Verfügbarkeit & Robustheit**
 - Systeme sind nicht zwangsweise Verfügbar
 - Verteilte Systeme unterliegen fehlerbehafteter Infrastruktur (z.B. Netzwerkausfall)

Heterogenität im Verantwortungsbereich

- **A2A – Application to Application**
 - Integration von Anwendungen und System innerhalb der selben Firma
- **B2B – Business to Business**
 - Die externe Integration von Geschäftspartner-, Kunden- und Lieferanten-Prozessen bzw. – Anwendungen
- **B2C – Business to Consumer**
 - Die direkte Integration von Endkunden in die internen Unternehmensprozesse