



Datenbanksysteme I

Vorlesungsskript

Mitschrift von Falk-Jonatan Strube

Vorlesung von Dr. Axel Toll

22. März 2016

Inhaltsverzeichnis

1 Datenbank als System und Modell	3
1.1 Daten als Unternehmensressource	3
1.1.1 Daten und Informationen	3
1.1.2 Klassifikation von Daten	4
1.1.3 Datenverschlüsselung	5
1.1.4 Speicher- und Zugriffsformen	7
1.2 Datenmodelle als Abbild	9
1.3 Datenbanksysteme als Grundlage	11
2 Datenbanksystem	12
2.1 Konventioneller / Datenbankorientierter Ansatz	12

Prüfungsmodalitäten

PVL unbenoteter Beleg als Voraussetzung zur Prüfung

- 1.) Access-Beleg (in Papier-Form abzugeben)
- 2.) Abnahme der SQL-Praktikums-Aufgaben (Abnahme während Praktikumszeit)

SP schriftliche Prüfung, 90min

keine eigenen Unterlagen zugelassen. Nur zuvor ausgegeben Referenzen.

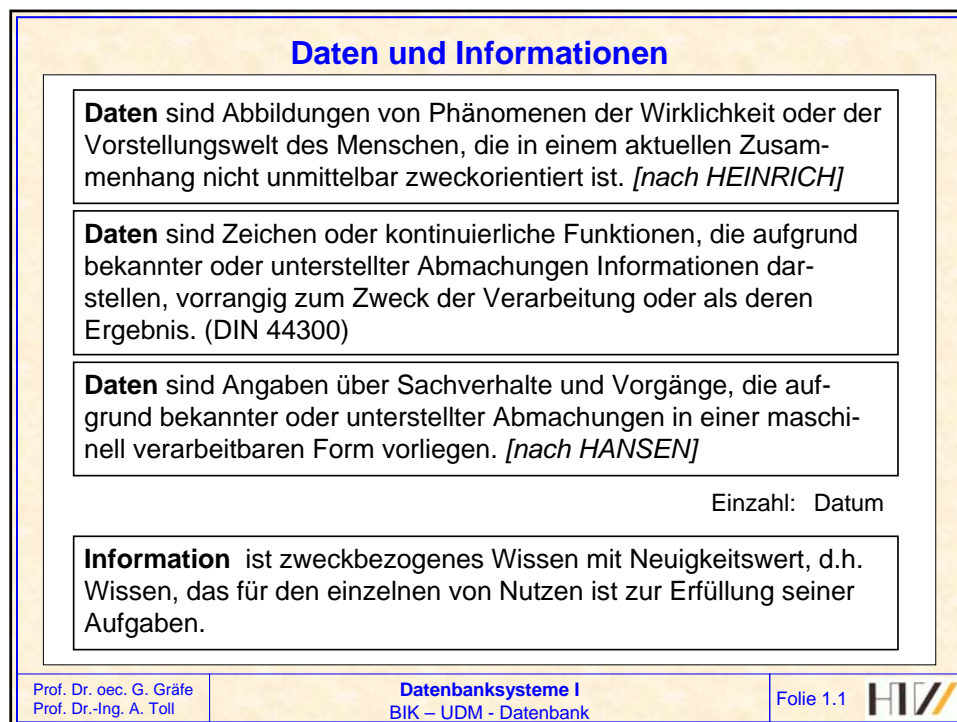
1 Betriebliche Informations- und Kommunikationssysteme - Unternehmensmodell - Datenbank

1.1 Daten als Unternehmensressource

1.1.1 Daten und Informationen

Redundante Daten bergen Gefahr von Inkonsistenz \Rightarrow Ziel: Schaffen von Datenbank mit folgenden Eigenschaften:


- ohne Inkonsistenzen (redundanzarm)
- Zugriffsschutz
- Mehrfachzugriff
- Backup-Möglichkeiten (mit Widerspruchsfreier Wiederherstellung)



	Daten	Informationen
Zweck	zweckneutral	zweckgebunden
Verarbeitung	maschinell	Interpretation durch Menschen
Speicherform	vergegenständlicht	an Menschen gebunden

Betriebliche Produktionsfaktoren

- klassische Faktoren
 - Betriebsmittel
 - Werkstoffe
 - Arbeitskraft
- Daten + Informationen

Probleme des Informatikeinsatzes		
➤	Softwarekrise trotz CASE Tools	mögliche Ursache: ungenügende Anforderungsanalyse und -definition
➤	Zu hoher Kostenaufwand in den letzten Entwicklungsphasen und in der Systemwartung	mögliche Ursache: unzureichende methodische Unterstützung der Anfangsphase
➤	Überschreitung von Lieferterminen	mögliche Ursache: mangelndes Projektmanagement
➤	Jahrhundertproblem der Informatik (Datenchaos)	mögliche Ursache: fehlende unternehmensweite Datenmodellierung
➤	Unzulängliche Anwendungssysteme, fehlende Nutzung	mögliche Ursache: unzureichende Einbeziehung der Nutzer in die Systementwicklung
<p>Das Jahrhundertproblem der Informatik besteht in:</p> <ul style="list-style-type: none"> ➤ Der Bewältigung des Datenchaos, das infolge unkontrolliert gewachsener Datenbestände fast überall entstanden ist. ➤ Der Schaffung einer einheitlichen, zentrale und dezentrale Datenbestände umfassenden Datenbasis, die für die effiziente Nutzung zukunftssträchtiger Möglichkeiten der Informatik - gemeint sind benutzerfreundliche, auch Nichtinformatikern zumutbare Anwendungs-generatoren und höhere Datenbanksprachen - unerlässlich ist. <i>[nach Vetter]</i> 		
Prof. Dr. oec. G. Gräfe Prof. Dr.-Ing. A. Toll	Datenbanksysteme I BIK – UDM - Datenbank	Folie 1.2 

Große Datenbestände ⇒ Maßnahmen zur Datenorganisation

Eine mögliche Organisationsform (logisches Konzept): Ablage in Relationen (=Tabelle)

Eine Zeile in dieser Tabelle nennt man *Datensatz* (Tupel, Record, ...).

Eine Spalte nennt man *Datenfeld*.

1.1.2 Klassifikation von Daten

Mögliche Kriterien für Datenfeld

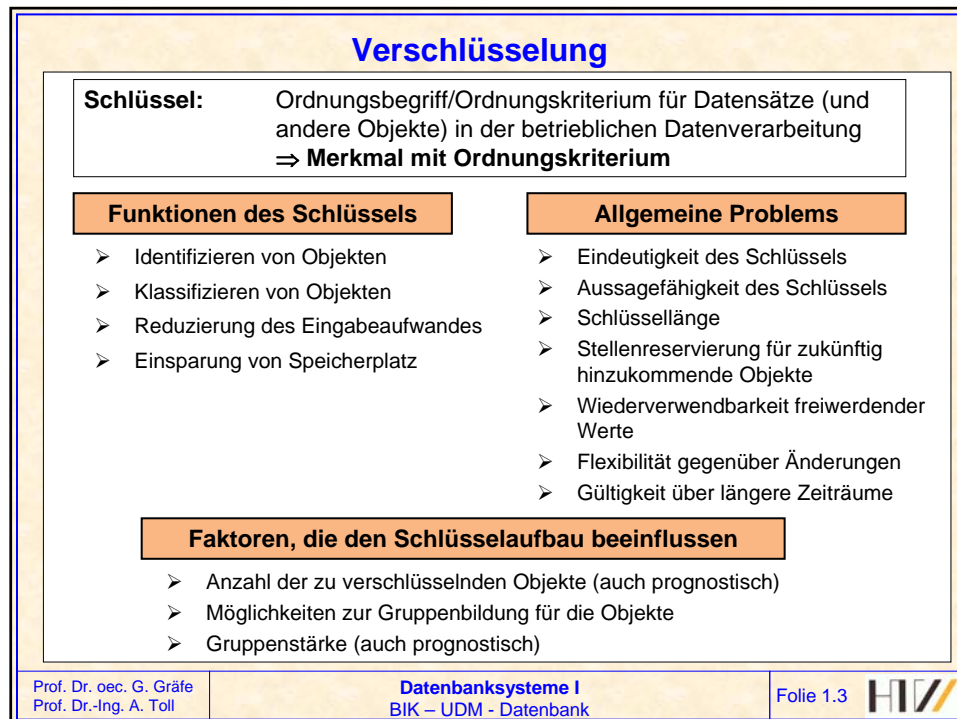
- Zeichenart
 - ganze Zahl ⇒ für Aufzählungen
 - reelle Zahl ⇒ numerische Berechnungen
 - Währung ⇒ finanztechnische Berechnungen
 - Datum ⇒ kalendarische Berechnungen/Werte
 - Text ⇒ Beschreibung
 - Bitmuster ⇒ Video, Bilder, ...
- Erscheinungsform

- sprachlich
- bildlich
- schriftlich
- Stellung im Verarbeitungsprozess (E - V - A)
 - Eingabe
 - Verarbeitung
 - Ausgabe
- Verarbeitbarkeit mittels IT
(Umwandlung in digitale Daten: analog → diskret → digital)
- Verwendungszweck

	Charakterisierung	Beispiel
Stammdaten	selten zu verändern (über längeren Zeitraum in Struktur und Inhalt konstant)	Personalstammdaten (Name, Adresse)
Änderungsdaten	Aktualisierung der Stammdaten	Änderung der Adresse
Bestandsdaten	Periodische Änderung des Wertes (Inhalt) von Feldern, Datenstruktur besteht über längeren Zeitraum konstant	Lagerbestände, Kassenbestände
Bewegungsdaten	Daten zur Aktualisierung des Wertes von Bestandsdaten	Lagerzugänge und -abgänge
Archivdaten	vergangenheitsbezogene Daten die über längeren Zeitraum aufbewahrt werden	Rechnungen, Buchungen der vergangenen 5 Jahre
Transferdaten	Daten, die von einem anderen Programm erzeugt wurden und an ein anderes transferiert werden	Verkauf von Kundenadressen
Vormerkdaten	Daten, die solange existieren, bis ein genau definiertes Ereignis eintritt	Reservierung einer Materialmenge im Lager

1.1.3 Datenverschlüsselung

Gemeint ist nicht die Codierung und Decodierung von Daten, sondern das Zuweisen von Schlüsseln zu Datensätzen.



Identifizierender Schlüssel

kennzeichnet Objekteindeutig

Bsp.:

- Personal-Nr.
- Material-Nr.

Klassifizierender Schlüssel

ordnet Objekt einer Klasse zu

Bsp.:

- Länderkennung: D, C, CH, ...
- Geschlecht: M, W

Hierarchischer Verbundschlüssel

identifizierender Teil hängt vom klassifizierenden Teil ab

Bsp.:

- Autokennzeichen: $\underbrace{DD}_{\text{klass.}} \underbrace{XY 715}_{\text{ident.}}$

Parallelschlüssel

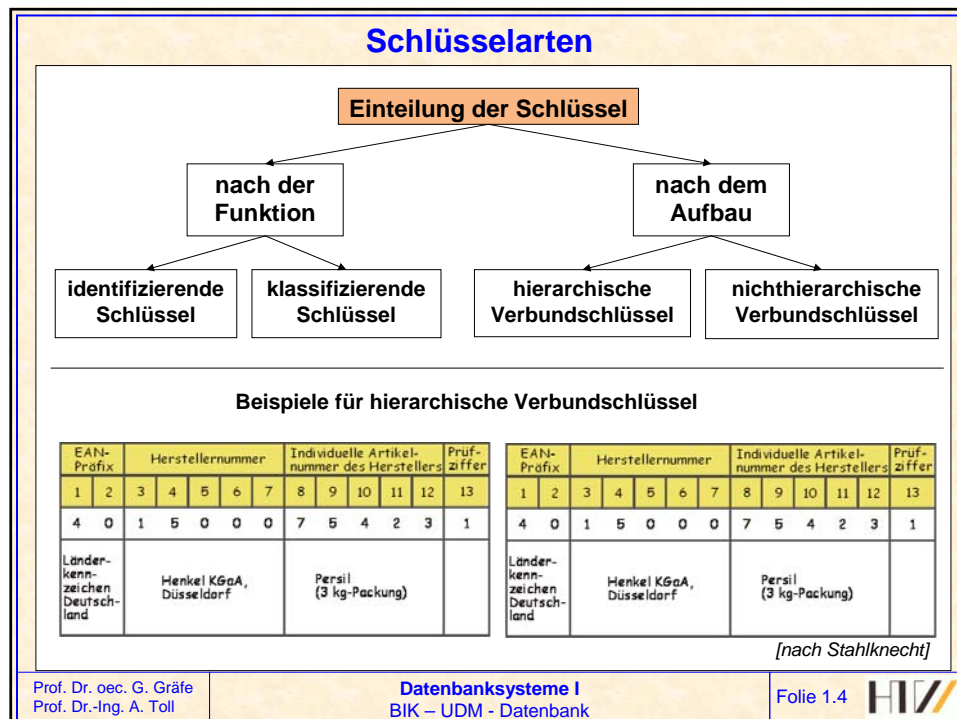
zwei unabhängige Schlüsselteile

Bsp.:

- Flugnummer $\underbrace{LH 283}_{\text{Flugnr.}} \underbrace{AB3}_{\text{Flugzeug}}$

spezielle Schlüssel in Datenbanksystemen

- **Primärschlüssel** (primary key PK): Datenfeld oder die Kombination aus Datenfeldern, die den Datensatz in der Tabelle eindeutig identifizieren.
Bsp. Vereinsdatenbank:
Primärschlüssel als einzelnes Datenfeld (Mitgliedertabelle): Mitglieds-ID
Primärschlüssel als eine Kombination von Datenfeldern (Beitragstabelle): ID mit Jahr (für Vereinsbeitrag abhängig von Jahr)
- **Fremdschlüssel** (foreign key FK): Datenfeld, oder Kombination aus Datenfeldern, der (die) auf den PK einer anderen Tabelle zeigt.
Bsp.: Mitglieds-ID in Tabelle mit Datenfelder-Primärschlüssel kommt aus der ersten Tabelle
- **Referentielle Integrität**: Jeder Wert eines FK muss gleich dem Wert des PK sein, auf den der FK zeigt.
Bsp.: Neuer Eintrag in Beitragstabelle kann nur neue Einträge bekommen, die Mitglieder aus Mitgliedertabelle enthält. Anders herum kann aus der Mitgliedertabelle kein Mitglied gelöscht werden, das noch in der Beitragstabelle genutzt wird.



1.1.4 Speicher- und Zugriffsformen


- **sequentielle Speicherung** (fortlaufend)
Bsp.: Bandlaufwerk

101	102	103	...
-----	-----	-----	-----
- **verkettete Speicherung**
Bsp.: verkettete Listen (vgl. Programmierung I)
- **indexverkettete Speicherung**
Trennung: Datenspeicherung und „Weg“ zu den Daten
 - Indexdatei (sortiert nach entsprechendem Index)

- ♦ Primärindex zeigt auf physische Adresse
- ♦ Sekundärindex zeigt auf Primärindex
- Hauptdatei

Grundprinzip der Indizierung				
	Mitnr	Name	Ort	Alter
ADR1	101	Hase	Dresden	37
ADR2	102	Igel	Dresden	19
ADR3	103	Fuchs	Dresden	23
ADR4	104	Elster	Freiberg	26
ADR5	105	Uhu	Berlin	22
ADR6	106	Rabe	Radebeul	68

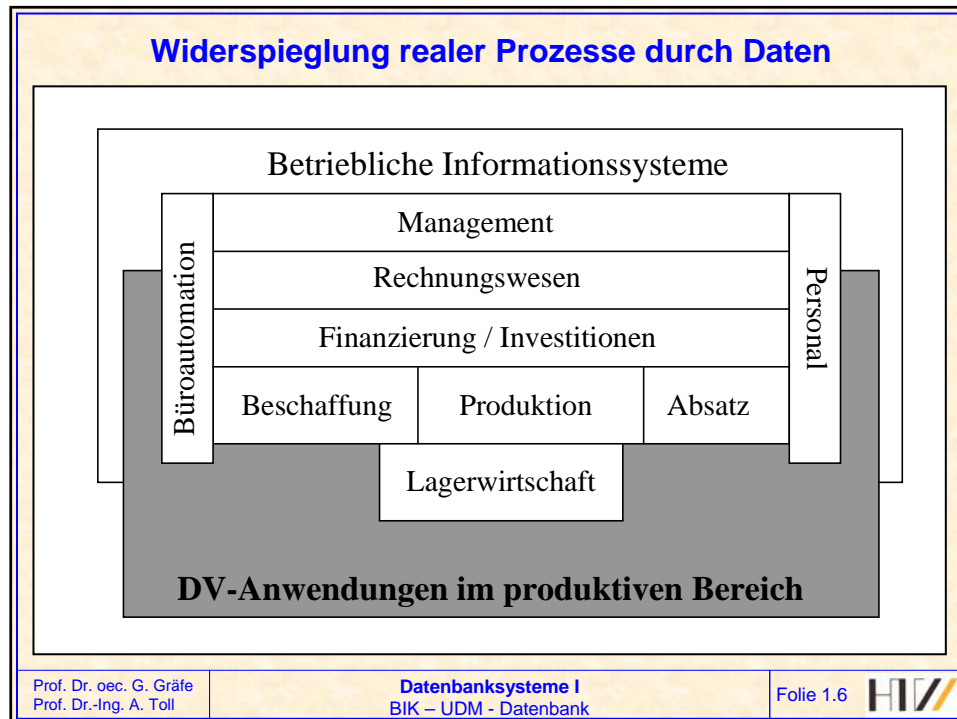
Primärindex Mitnr		Sekundärindex Ort		Sekundärindex Alter	
Feldinhalt	Adresse	Feldinhalt	Schlüssel	Feldinhalt	Schlüssel
101	ADR1	Berlin	105	19	102
102	ADR2	Dresden	101	22	105
103	ADR3	Dresden	102	23	103
104	ADR4	Dresden	103	26	104
105	ADR5	Freiberg	104	37	101
106	ADR6	Radebeul	106	68	106

Prof. Dr. oec. G. Gräfe	Datenbanksysteme I	Folie 1.5
Prof. Dr.-Ing. A. Toll	BIK – UDM - Datenbank	

Unterschied Primärschlüssel-Primärindex:

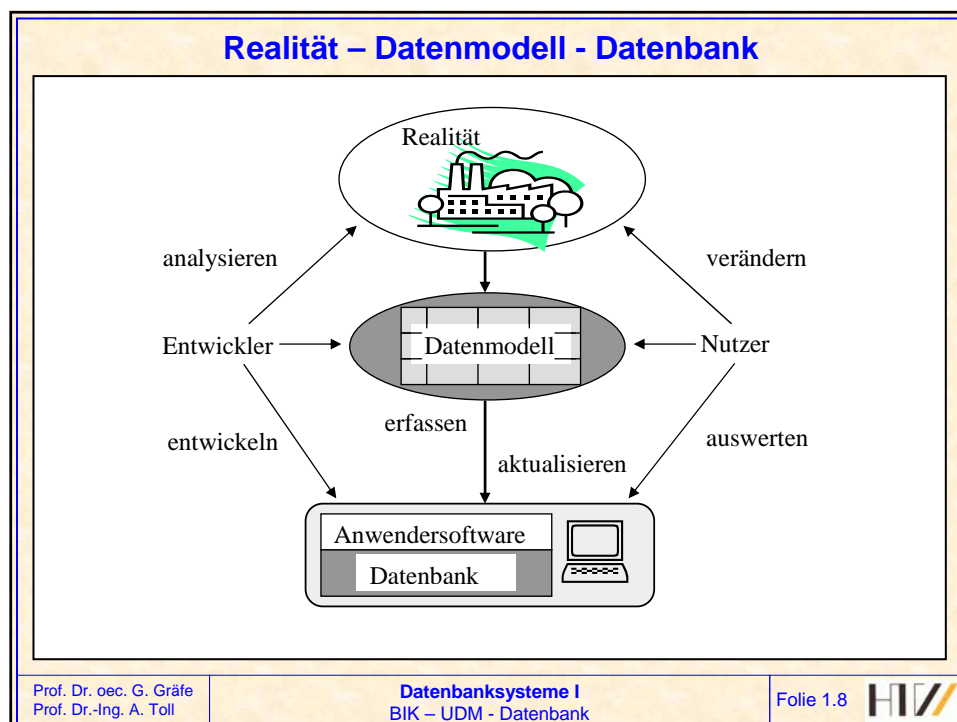
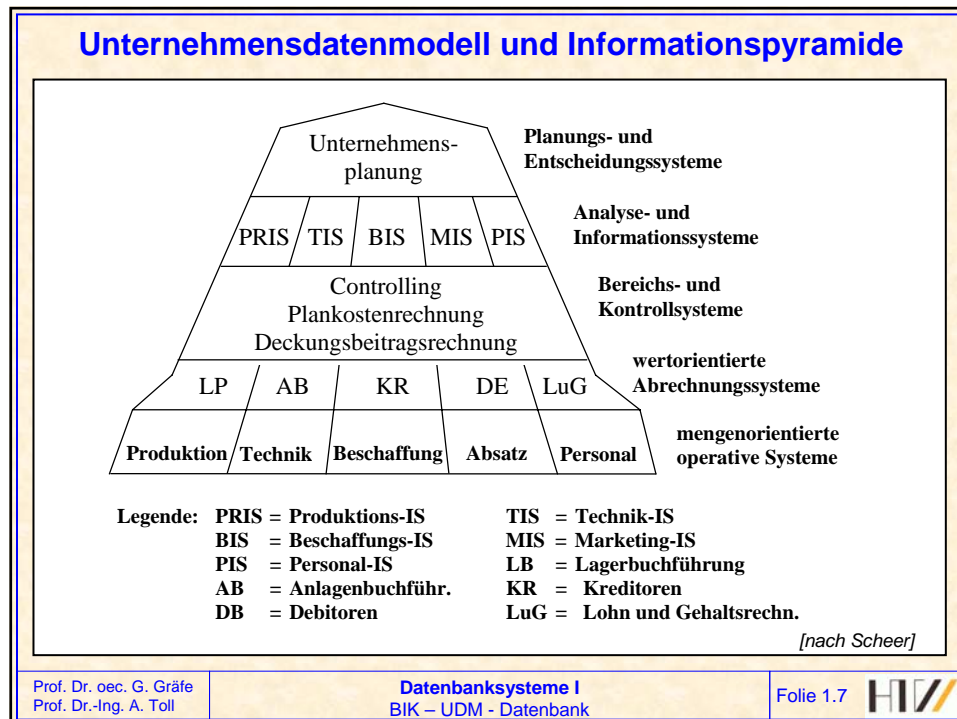
- Primärschlüssel dient dem Identifizieren
- Primärindex zum schnellen Suchen


1.2 Datenmodelle als informationelles Abbild der Unternehmensrealität



Informationssystem

- *Funktionsmodell* (was soll das System leisten: Produktion, Lager, Beschaffung, ...) ⇒ Kernfrage: „Was will ich machen“
Strukturen, Abläufe
Technik: Programm-Ablauf-Plan (PAP), Ereignisorientierte Prozessketten (EPK), ...
- *Datenmodell*
Daten und deren logische Struktur
Technik: Entity-Relationship-Modell (ERM)



Datenmodellierung		
<p><u>Datenmodellierung:</u></p> <p>Aufbau einer einheitlichen, anwendungsübergreifenden Sicht der Datenressourcen des Unternehmens und Abbildung von Informationsprozessen und -beziehungen in die konkrete Struktur eines Datenverwaltungs- oder Datenbanksystems</p> <p>Die Datenmodellierung schafft die Voraussetzung für systematische Integration der Daten und damit für eine unternehmensweite Nutzung der Ressource Information.</p>		
Prof. Dr. oec. G. Gräfe Prof. Dr.-Ing. A. Toll	Datenbanksysteme I BIK – UDM - Datenbank	Folie 1.9 

Bsp.:

ABB9 (1-3)

1.3 Datenbanksysteme als technologische Grundlage der Datenverwaltung


Grundbegriffe der „Datenbankwelt“		
<p>Datenbank (Datenbasis) enthält die nach einheitlichen Gesichtspunkten gespeicherte Menge der Objekte, Objektklassen und Beziehungen eines Bereiches (einschließlich zugehöriger Organisations- oder Metadaten)</p>		
<p>Datenbankbetriebssystem (Database-Managementsystem –DBMS) Software, die es aufbauend auf dem Rechnerbetriebssystem gestattet, die Datenmenge effektiv und rationell zu speichern, zu pflegen, auszuwerten und Sonderfunktionen, wie Datensicherung, Zugriffsschutz u.a. auszuführen.</p>		
<p>Datenbanksystem Datenbank (Datenbanken) und Datenbankbetriebssystem bilden zusammen das Datenbanksystem (Datenbasissystem)</p>		
Prof. Dr. oec. G. Gräfe Prof. Dr.-Ing. A. Toll	Datenbanksysteme I BIK – UDM - Datenbank	Folie 1.10 

ABB10

Datenbasis: Tabellen mit Metadaten

Datenbankbetriebssystem (DBMS): Software, die mit Datenbasis kommuniziert

2 Grundlagen und Architektur eines Datenbanksystems (DBS)

2.1 Defekte des konventionellen Ansatzes der Datenverwaltung / Zielstellung des datenbankorientierten Ansatzes

konventionell

ABB 11

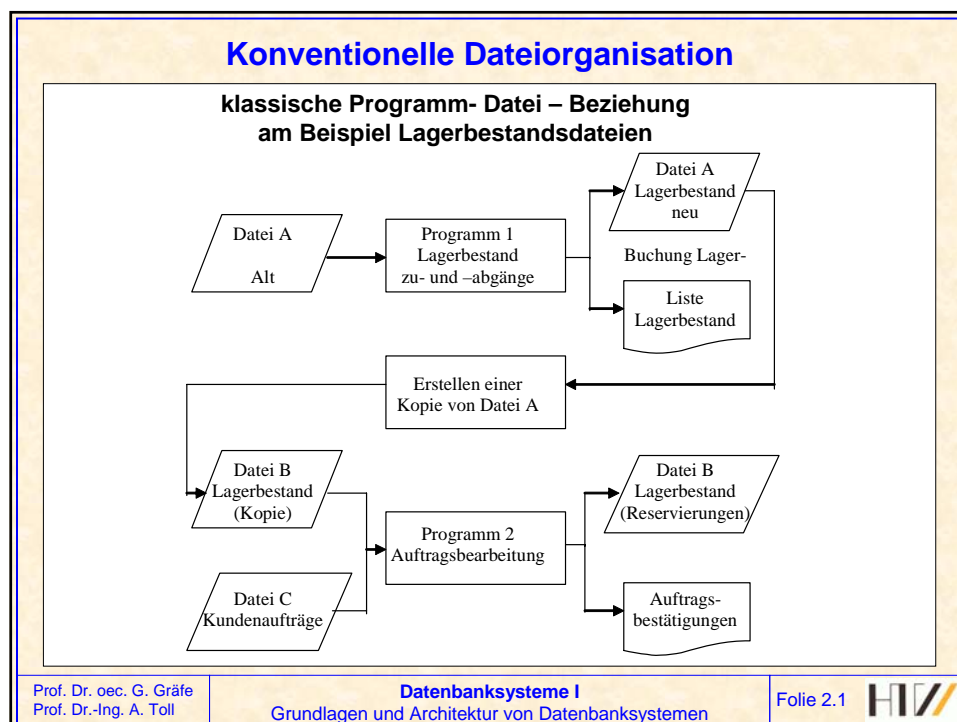
konventionelle Datenorganisation

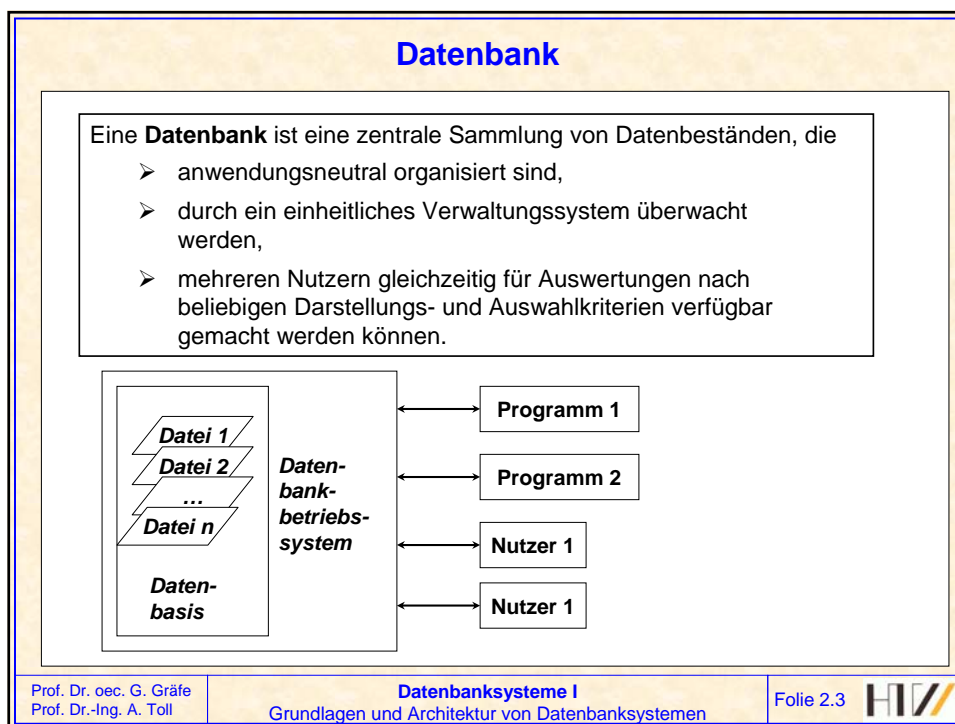
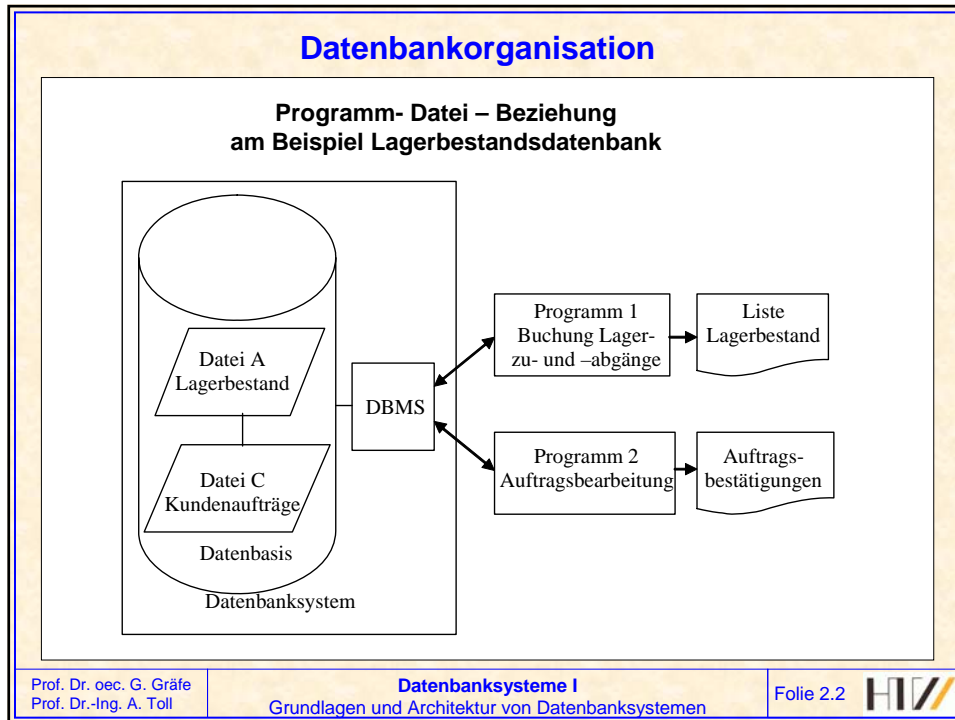
Merkmale

- Datenspeicherung je Anwendung
- Datenspeicherung auf physischem Niveau

Nachteile

- mangelnde Passfähigkeit (Zugriffskonflikte usw.)
- Redundanz
- Konsistenzprobleme
- mangelnde Flexibilität
- Daten-Programm-Abhängigkeit (kurz: Datenabhängigkeit)





Zielsetzung des Datenbankeinsatzes

Zielstellungen des Einsatzes von Datenbanken


1. Reduzierung der Redundanz auf ein zweckmäßiges Maß (Redundanzarmut)
2. Sicherung der Integrität und Konsistenz der Daten (Datensicherheit)
3. Erreichung einer Datenunabhängigkeit der Programme
4. Integrierte und flexible Auswertbarkeit der Daten
5. Gewährleistung eines hinreichenden Zugriffsschutzes
6. Bedienung EDV-ferner Nutzer und Erreichung eines nutzerfreundlichen Dialogs

Integrität bedeutet ganz allgemein inhaltliche Korrektheit/Richtigkeit und Vollständigkeit der Daten.

Konsistenz bedeutet logische Übereinstimmung und innere Widerspruchsfreiheit der Daten.

Prof. Dr. oec. G. Gräfe
Prof. Dr.-Ing. A. Toll

Datenbanksysteme I
Grundlagen und Architektur von Datenbanksystemen

Folie 2.4 

1.) Bsp. für gewollte Redundanz: Sekundärindex

2.) Datensicherheit:

- physisch, falls bspw. der Server abbrennt
- logisch, dass bspw. alle Daten den richtigen Typ haben