Datenbanken

Andreas Heß
Hochschule Furtwangen

Zur Person

- Andreas Heß, 39, aus Mainz
- Kurzlebenslauf:
  - Dipl.-Inform. (FH) 2001, FH Darmstadt
  - Ph.D. 2006, University College Dublin
  - Vrije Universiteit Amsterda
  - Lycos Europe
  - Deutsche Nationalbibliothek

Wo kommen da Datenbanken vor?

Administratives

- Leistungsnachweise: Praktika und Klausur
- Für WIB:
  Vorlesung und Praktikum getrennte Fächer
- Als WPV:
  nur ein Fach mit Vorlesung und Praktikum
- Klausur vs. 20. Juli, 15:45 Uhr, Aula

WIB

49 Einträge

Datenbanken Praktikum WIB1 SS17

Praktikum zur Vorlesung Datenbanken

Sommersemester 2017
I2.20

(0)

Infoseite | starten

Datenbanken WIB 1 (Heß)

Schichtenaufbau von DBMS, Modellierung mit ERM, Anfragesprachen, Transaktionen usw.

(0)

Infoseite | starten

Praktika

- Praktikum beginnt ab Mi., 29. März in I2.20
  - Gruppe 1: 14:00–15:30 in I2.20
  - Gruppe 2: 15:45–17:15 in I2.20
- Bearbeitung der Aufgaben in Zweiergruppen
- Am besten eigenen Laptop mitbringen!

Praktika

- Aufgaben werden jeweils Mi. nach der Vorlesung ausgegeben
- späteste Abgabe:
  Mi. eine Woche später, 23:59 Uhr

  evtl. Ausnahmen werden gesondert bekanntgegeben

Praktika

- Einreichung als **PDF** über FELIX
- sowie **SQL-Dump** der Datenbank
- in einem **ZIP**-File
- Mindestens ein Mitglied jeder Gruppe muss zur Besprechung ins Tutorium kommen

Praktika – Bewertung

- „Ampel“-Bewertung pro Aufgabe:
  - Komplett richtig → grün → 2 Punkte
  - Ansatz richtig → gelb → 1 Punkt
  - Komplett falsch* → rot → 0 Punkte

* oder nicht fristgerecht eingereicht

Praktika – Bewertung

- Mindestens 50% der Punkte zum Bestehen des Praktikums nötig

Willing is not enough; we must do.
Knowing is not enough; we must apply.

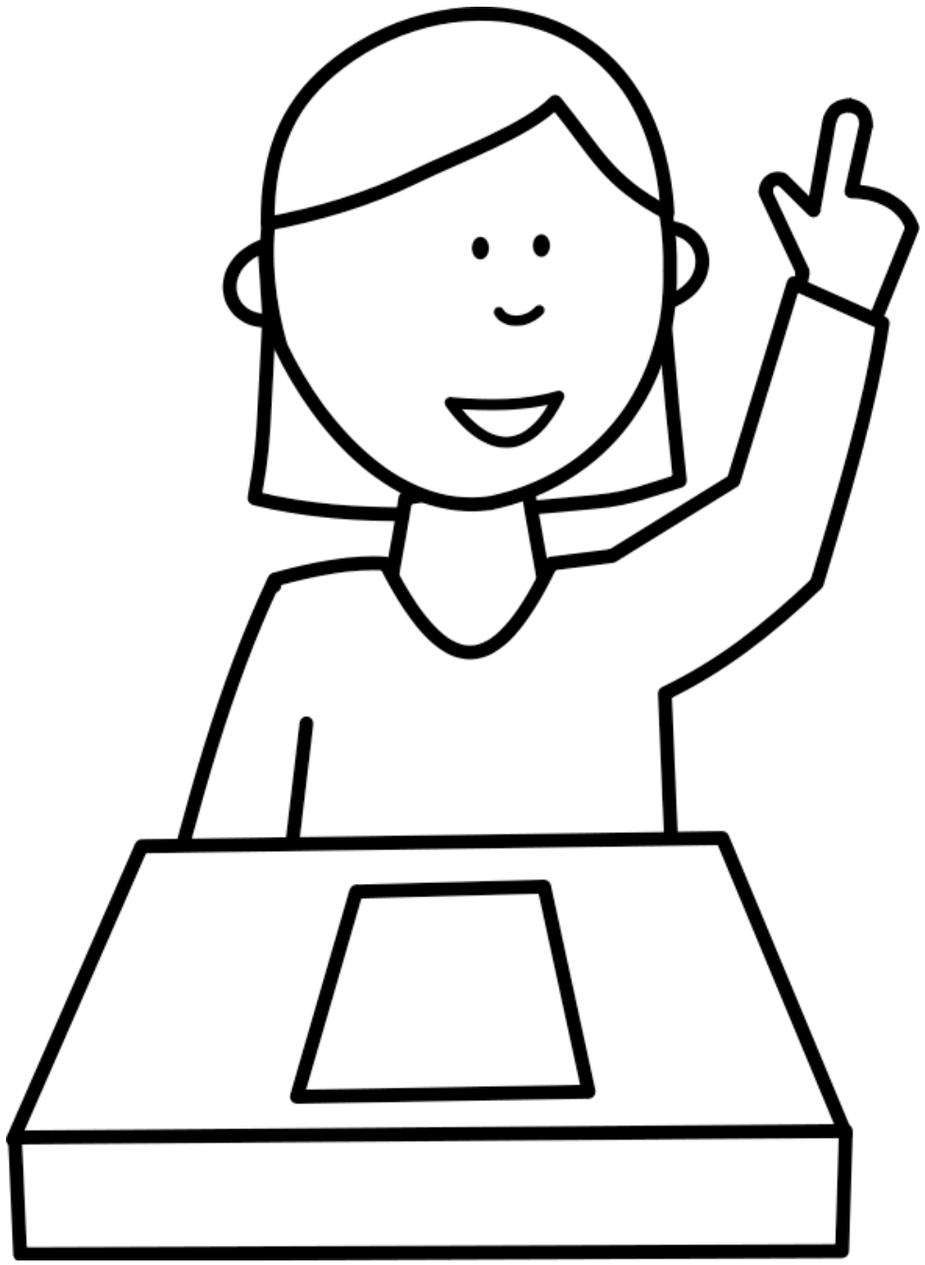
李小龍
(Bruce Lee)

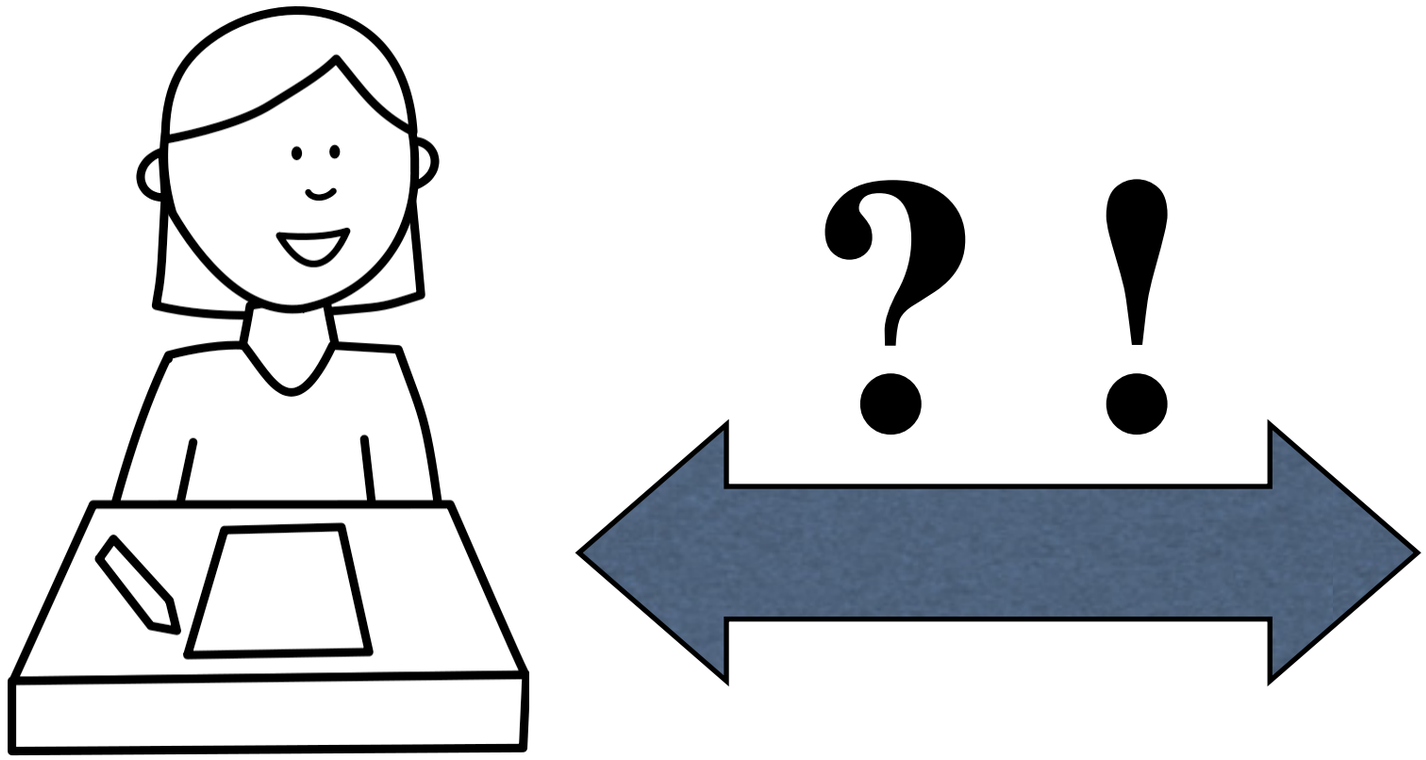
Erstes Praktikum

- Aufgabe für das erste Praktikum:
  Installieren Sie eine Datenbank-Software;
  z.B.:

- Im Tutorium wird gezeigt, wie‘s geht

Vorlesung





hea@hs-furtwangen.de

I 2.08

Sprechzeiten
Dienstag 11:30 Uhr
oder nach
Vereinbarung

Literaturempfehlung

- Kemper, Alfons und Eickler, André:
  Datenbanksysteme - Eine Einführung
- Piepmeyer, Lothar:
  Grundkurs Datenbanksysteme
- Eine SQL-Referenz nach Wahl

Datenbank

- Was ist das?
- Wofür braucht man das?

Anwendungen

- Beispiele:
  - E-Commerce
  - Web-Anwendungen
  - Versicherungen
  - Bibliotheken

- Welche Daten werden gespeichert?

Adressverwaltung

- Beispiel: Adressverwaltung
- Wieso braucht man nun eine Datenbank?

DATABASES

Y U NO USE SIMPLE TEXT FILE?

In einer Welt ohne Datenbanken…

Adressdaten (0)

```
Kemper, Alfons, Schillerstr. 15, 55116 Mainz
Hans Meier, Goethestraße 23a, 12345 Mönchengladbach
Andreas Heß; Lessingstr. 25; 78120 Furtwangen
Albrecht, Alexandra  Marktplatz 1  69117 Heidelberg
Dampf, Hans                Alle Gassen 2 78120 Furtwangen
```

Muss man was beachten, wenn man diese Daten maschinell verarbeiten will?

Adressdaten (1)

Name
Vorname
Straße
Postleitzahl
Ort
E-Mail-Adresse

Muss man noch was beachten, wenn man diese Daten maschinell verarbeiten will?

Adressdaten (2)

Name	`20 Zeichen`
Vorname	`30 Zeichen`
Straße	
Postleitzahl	
Ort	`20 Zeichen`
E-Mail-Adresse	`20 Zeichen`

Warum diese Festlegung?

Karteikarten

- Durchsuchen usw. viel einfacher, wenn Karteikarten gleich aussehen

Adressdaten in Datei

Kemper	Alfons	Schillerstraße 15	55116	Mainz	kemper@example.com
Meier	Hans	Goethestraße 23a	12345	Mönchengladbach	hans@gmx.cz
Heß	Andreas	Lessingstr. 25	78120	Furtwangen	hea@hs-furtwangen.de
Albrecht	Alexandra	Marktplatz 1	67543	Heidelberg	alex@albrecht.com

Adressdaten in Datei

Kemper			116	Mainz	kemper@example.com
Meier	Hans	Goethestraße 23a	12345	Mönchengladbach	hans@gmx.cz
Heß	Andreas	Lessingst			angen.de
Albrecht	Alexandra	Marktplat			t.com

Zugriff auf Adresse von Hans Meier - wie?

Zeit- oder Programmieraufwand!
Vorsortieren?
Alle Daten im Arbeitsspeicher halten?

Adressdaten in Datei

Kemper	Alfons	Schillerstraße 15	55116	Mainz	kemper@example.com
Me			12345	Mönchengladbach	hans@gmx.cz
Heß	Andreas	Lessingstr. 25	78120	Furtwangen	hea@hs-furtwangen.de
Albrecht	Alexandra	Marktplatz 1			

Postleitzahl falsch

Konsistenz/Plausibilität muss gesichert sein

Aber in eine Datei kann ich alles mögliche rein schreiben

Beobachtungen (1)

- Bisher erkannte Nachteile:
  - Grundfunktionen erfordern Programmieraufwand
    → Standardisierung wünschenswert
  - Programmierer muss sich selbst um Konsistenz und Plausibilitätsprüfung kümmern, kein automatischer Schutz

Adressdaten in Datei

Kemper	Alfons	Schillerstraße 15	55116	Mainz	kemper@example.com
Meier	Hans	Goethestraße 23a	12345	Mönchengladbach	hans@gmx.cz
Heß	Andreas	Lessingstr. 25	78120	Furtwangen	hea@hs-furtwangen.de
Albrecht	Alexandra	Marktplatz 1	69117	Heidelberg	alex@albrecht.com
Dampf	Hans	Alle Gassen 17	78120	Furtwangen	hans@dampf.de

→ Manche Informationen doppelt

Redundanz

Adressdaten in Datei

Kemper				Mainz	kemper@example.com
Meier				Mönchengladbach	hans@gmx.cz
Heß	Andreas	Lessingstr. 25	78120	Furtwangen	hea@hs-furtwangen.de
Heß	Andreas	Goethestr. 39	60318	Frankfurt	hea@hs-furtwangen.de
Albrecht	Alex		67543	Heidelberg	alex@albrecht.com

Eine Person mit mehreren Adressen - wie?

Redundanz

➔ Manche Informationen doppelt

➔ unklar, ob es sich um die selbe Person handelt

Beobachtungen (2)

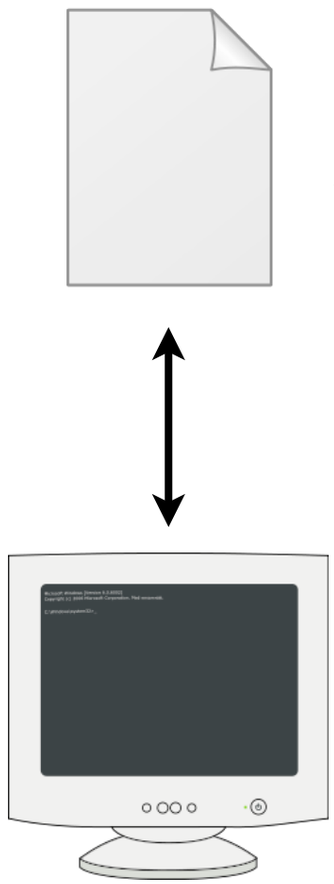
- Weitere Beobachtungen:
  - Sinnvolle Datenhaltung erfordert durchdachte Modellierung
    → Unterstützung durch Software wünschenswert
  - Redundanzfreiheit nicht sichergestellt

Beobachtungen (2)

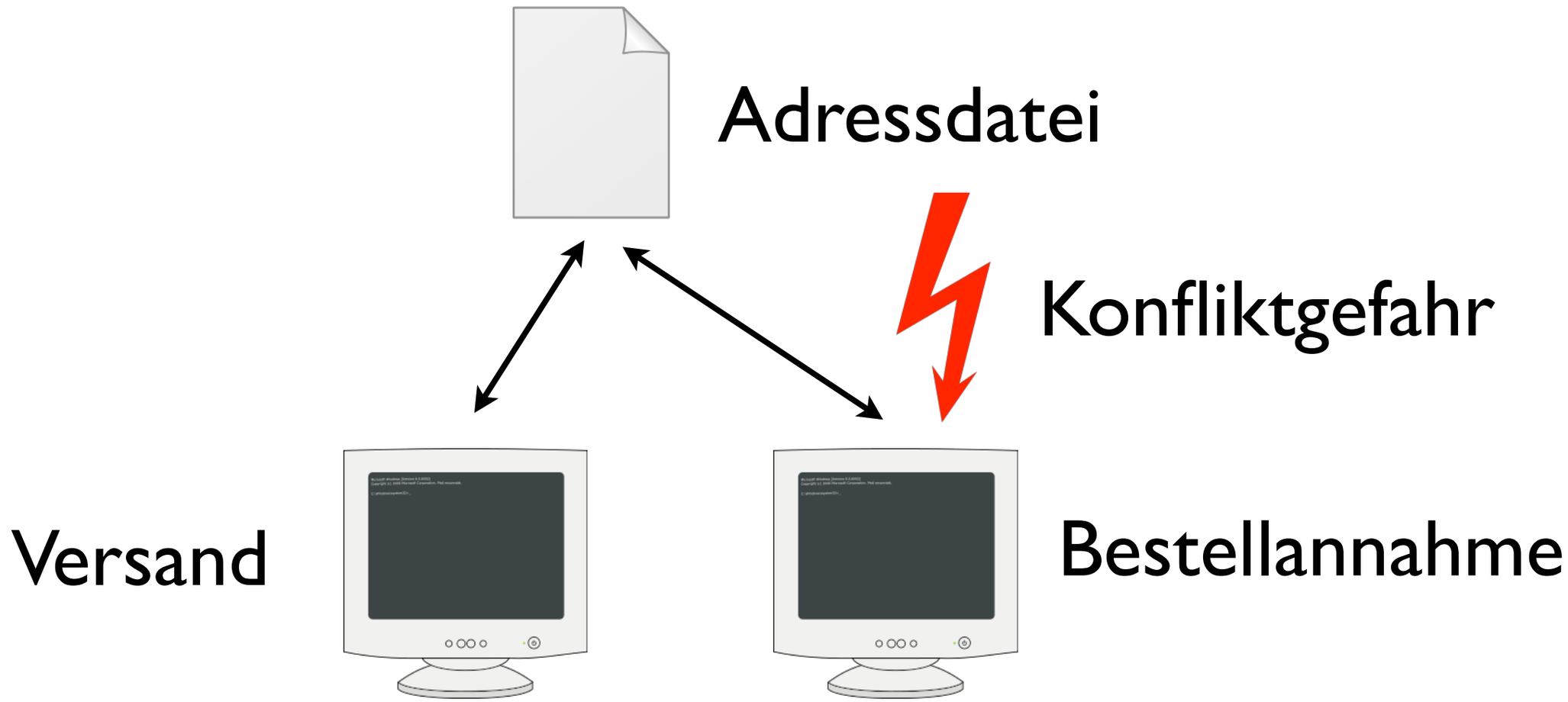
- Weitere Beobachtungen:
  - Sinnvolle Datenhaltung erfordert durchdachte Modellierung
    → Unterstützung durch Software wünschenswert
- Redundanzfreiheit nicht sichergestellt

Gibt es auch erwünschte Redundanz?

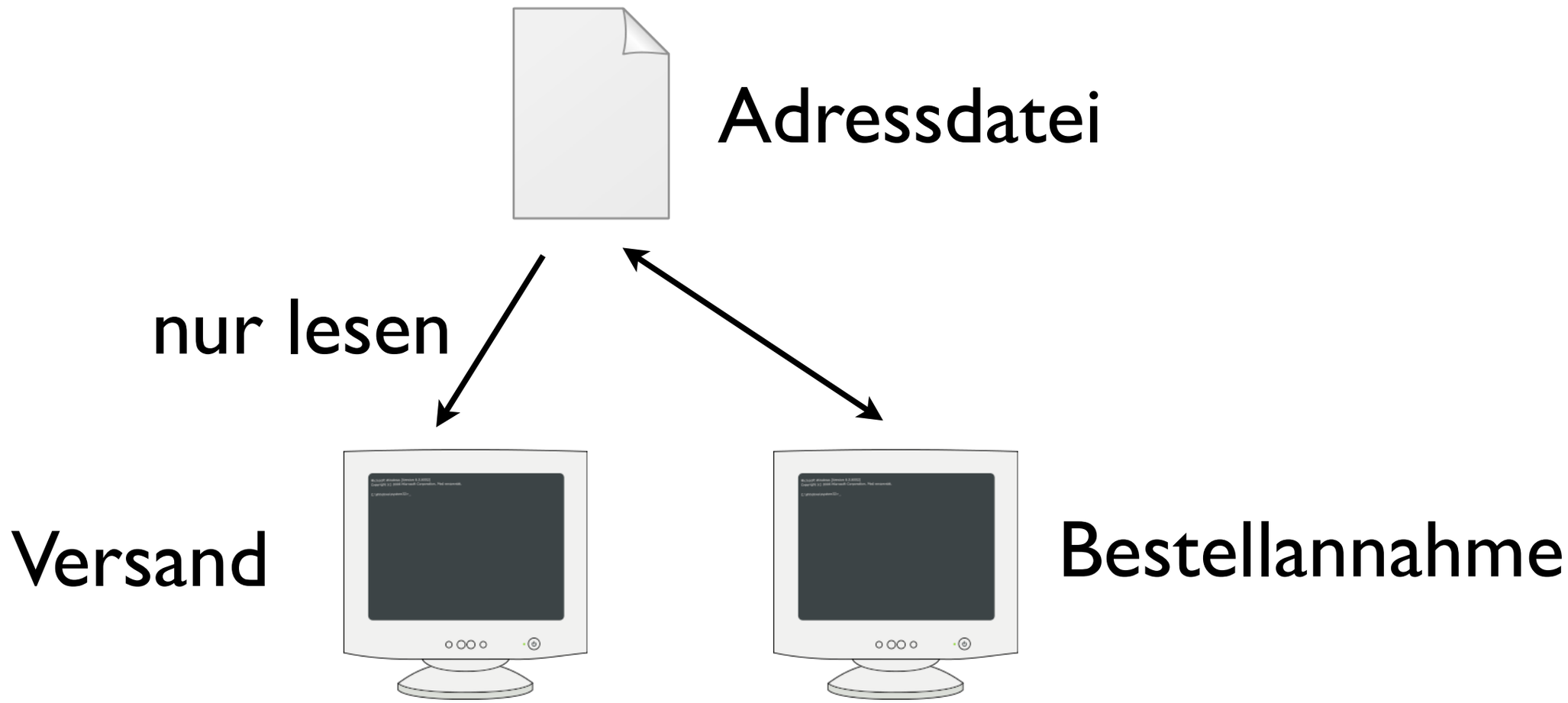
Zugriff



Zugriff

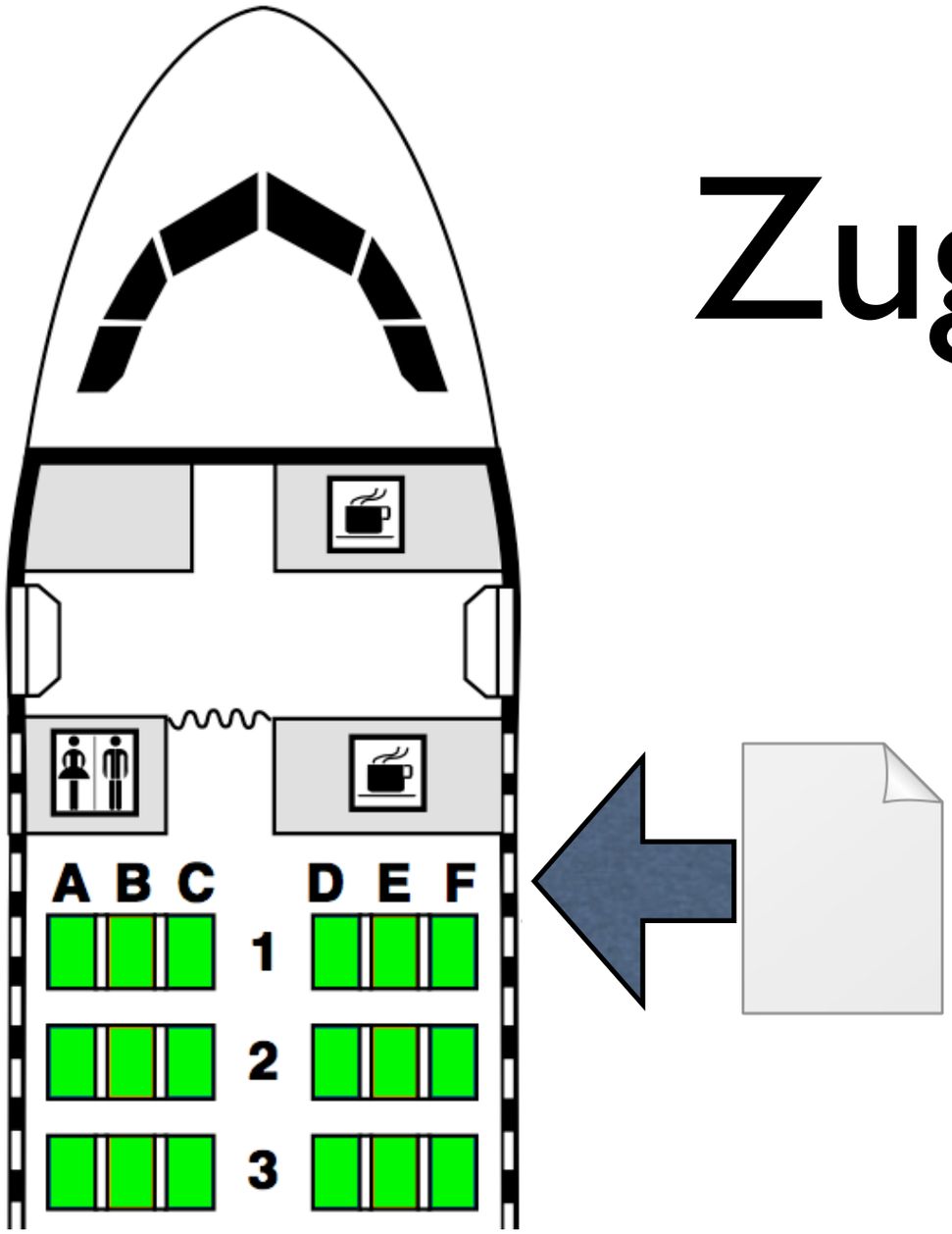


Zugriff



Problem gelöst?

Zugriffskonflikt

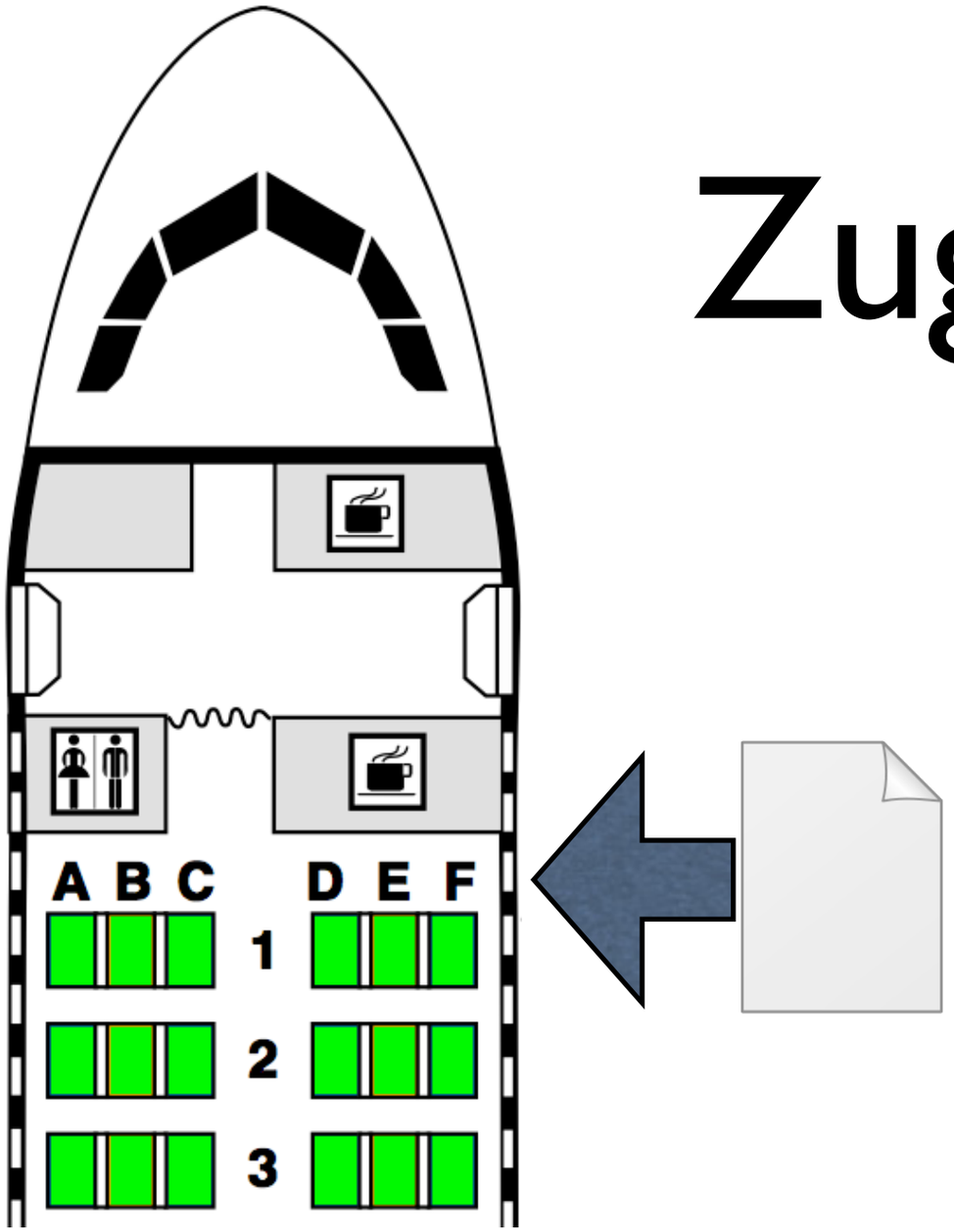


Reservierungsdatei

Flugbörse

Reisebüro Meyer

Zugriffskonflikt



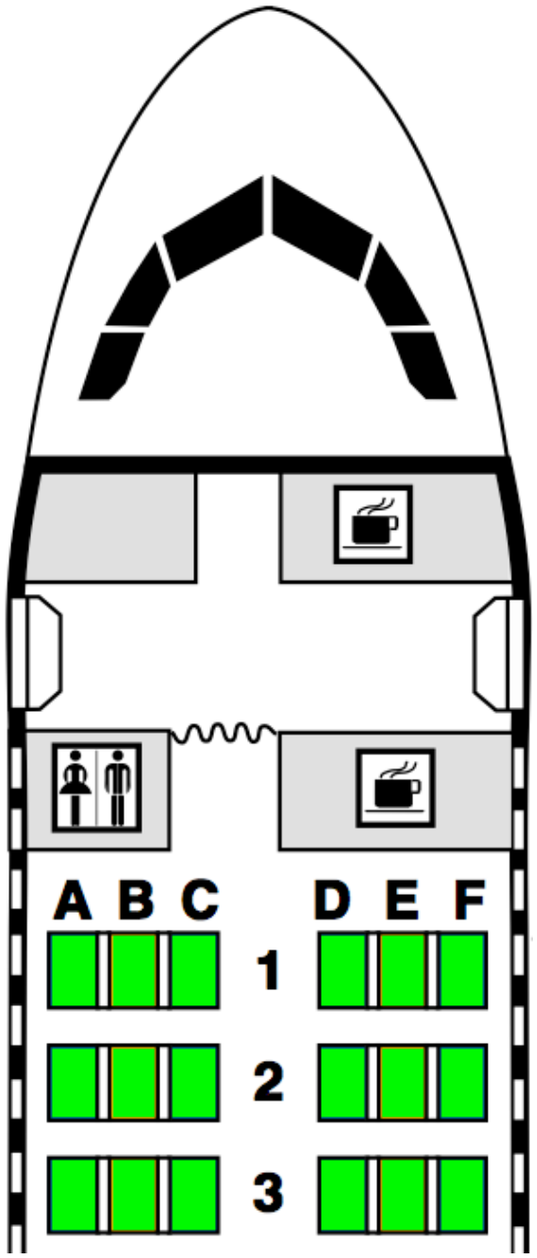
Reservierungsdatei

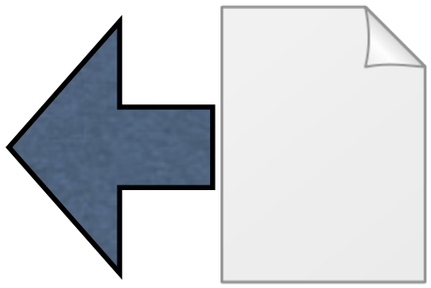
Flugbörse

Reisebüro Meyer

Platz 2D frei?

Zugriffskonflikt





Reservierungsdatei

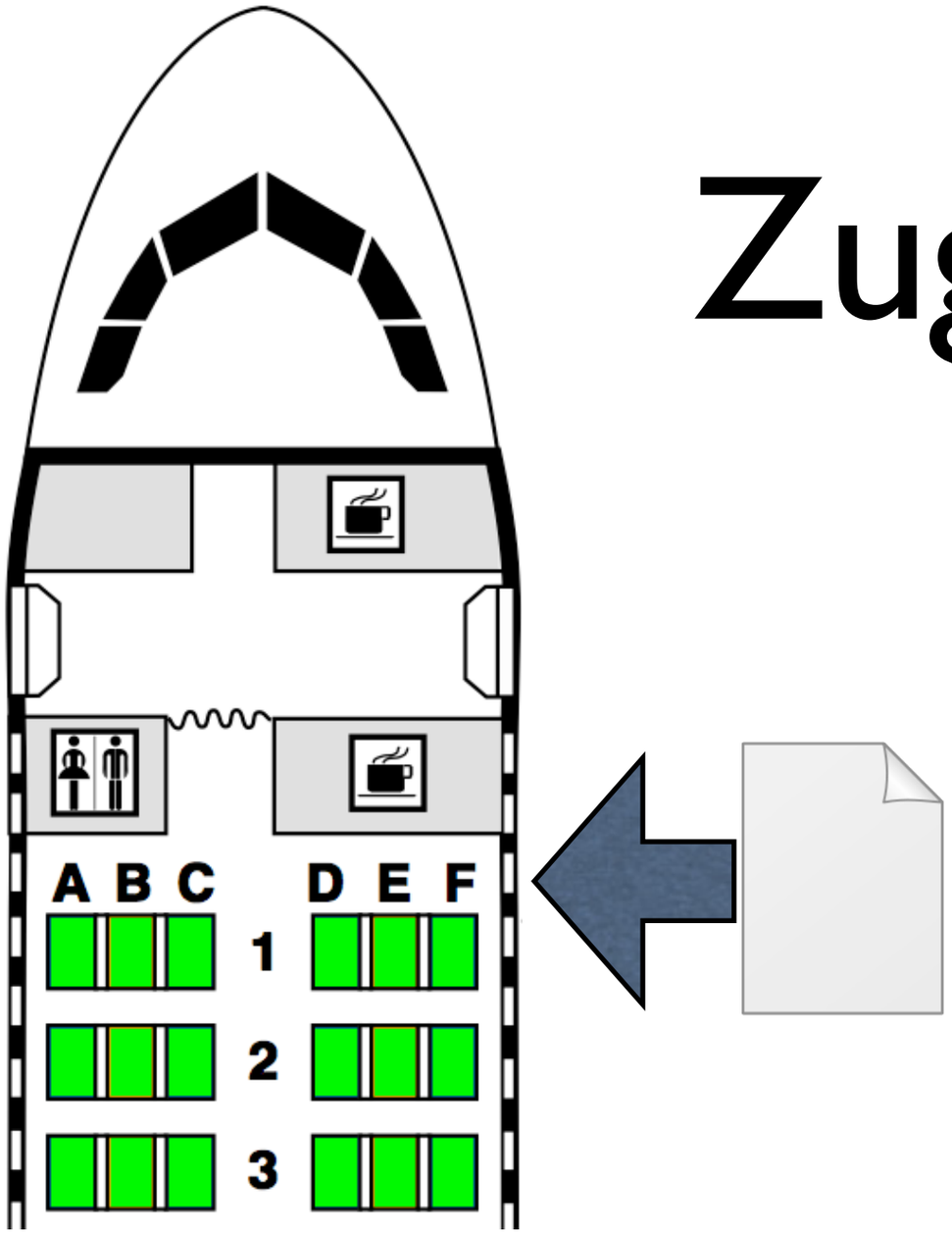
lesen

Flugbörse

Reisebüro Meyer

Platz 2D frei?

Zugriffskonflikt



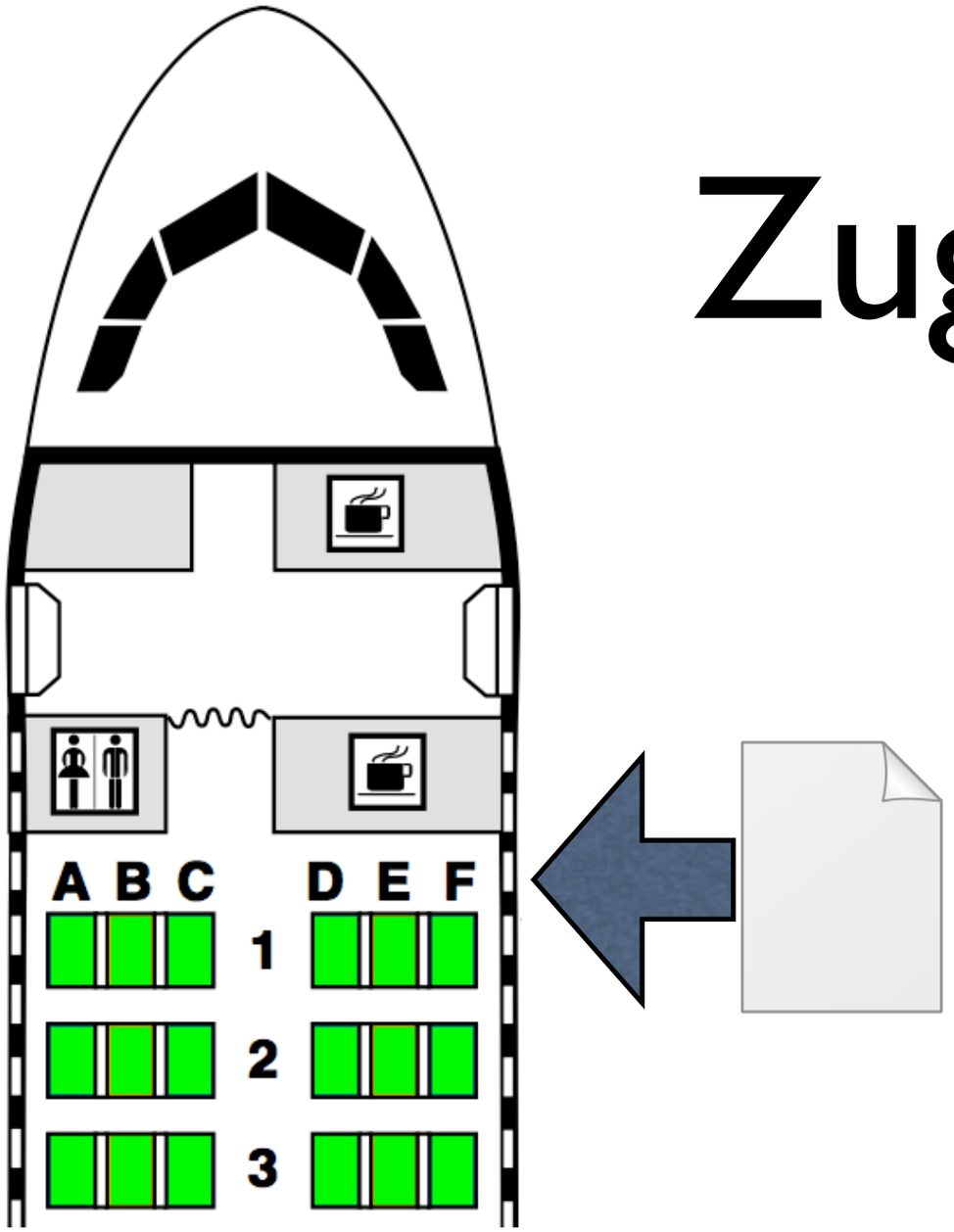
Reservierungsdatei

Flugbörse

Reisebüro Meyer

Platz 2D frei!

Zugriffskonflikt



Reservierungsdatei

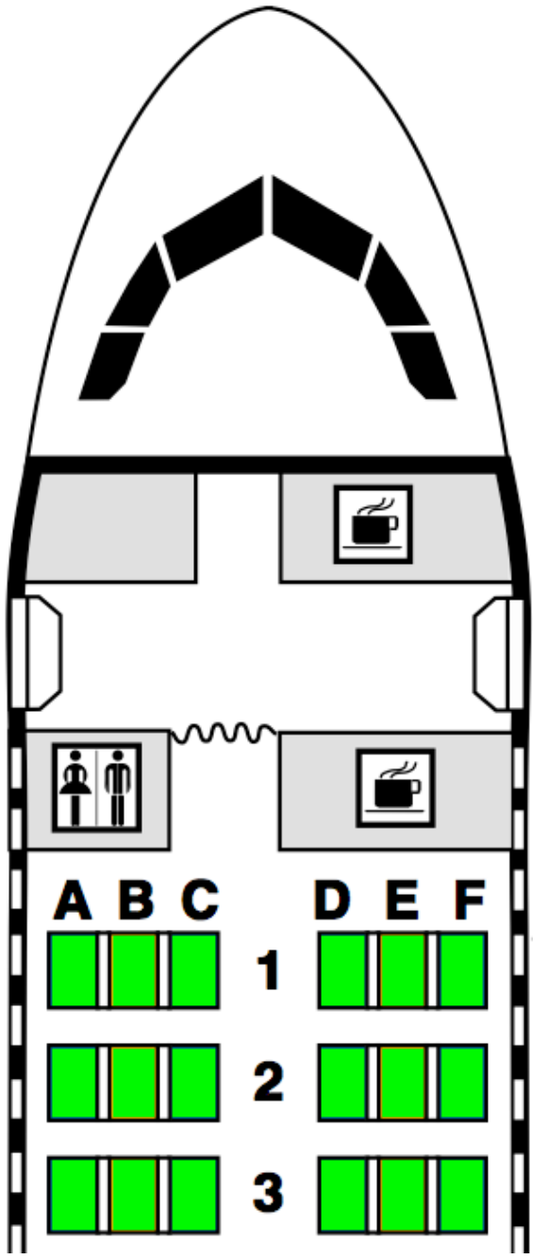
Flugbörse

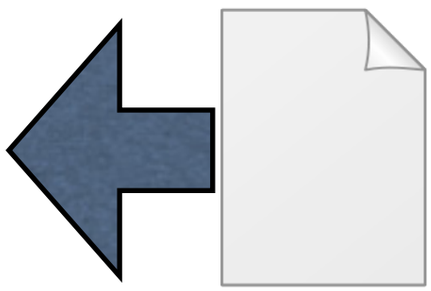
Reisebüro Meyer

Platz 2D frei!

Platz 2D frei?

Zugriffskonflikt





Reservierungsdatei

lesen

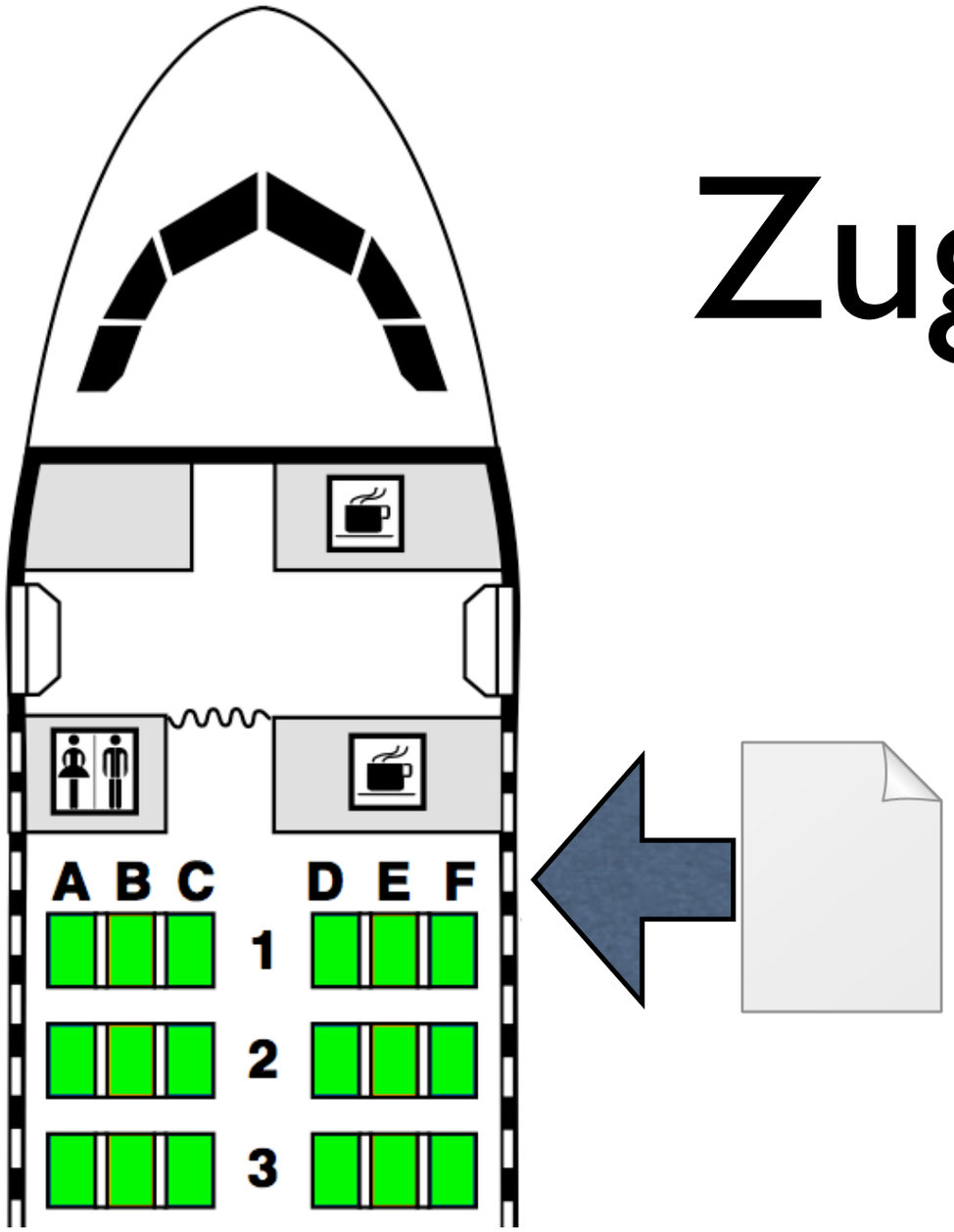
Flugbörse

Reisebüro Meyer

Platz 2D frei!

Platz 2D frei?

Zugriffskonflikt



Reservierungsdatei

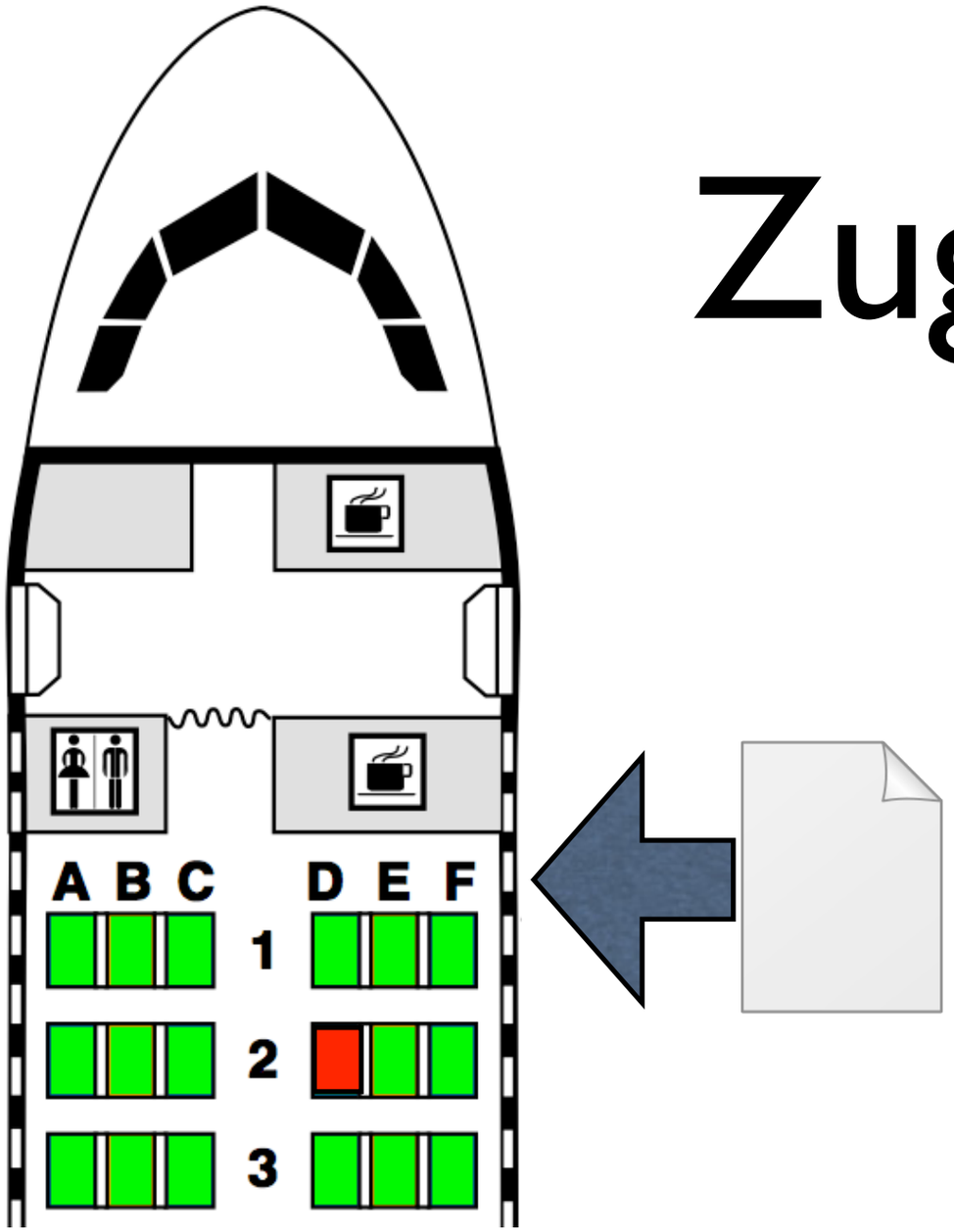
Flugbörse

Reisebüro Meyer

Platz 2D frei!

Platz 2D frei!

Zugriffskonflikt



Reservierungsdatei

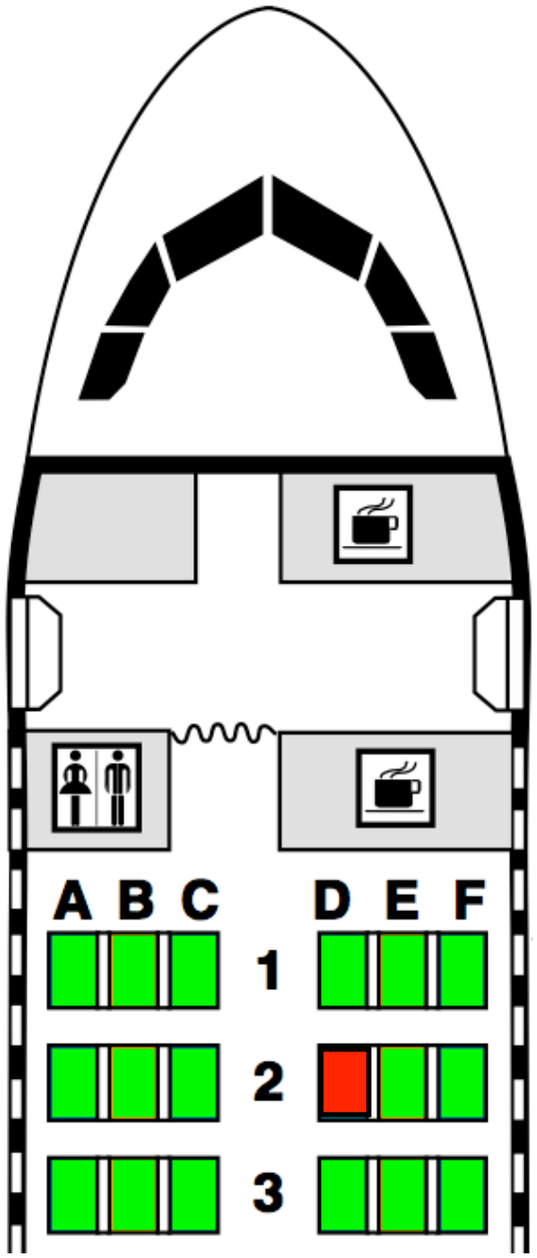
Flugbörse

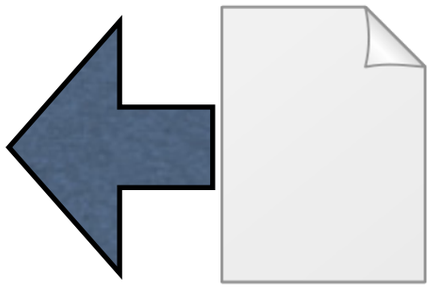
Reisebüro Meyer

Platz 2D: Alexandra Albrecht

Platz 2D frei!

Zugriffskonflikt





Reservierungsdatei

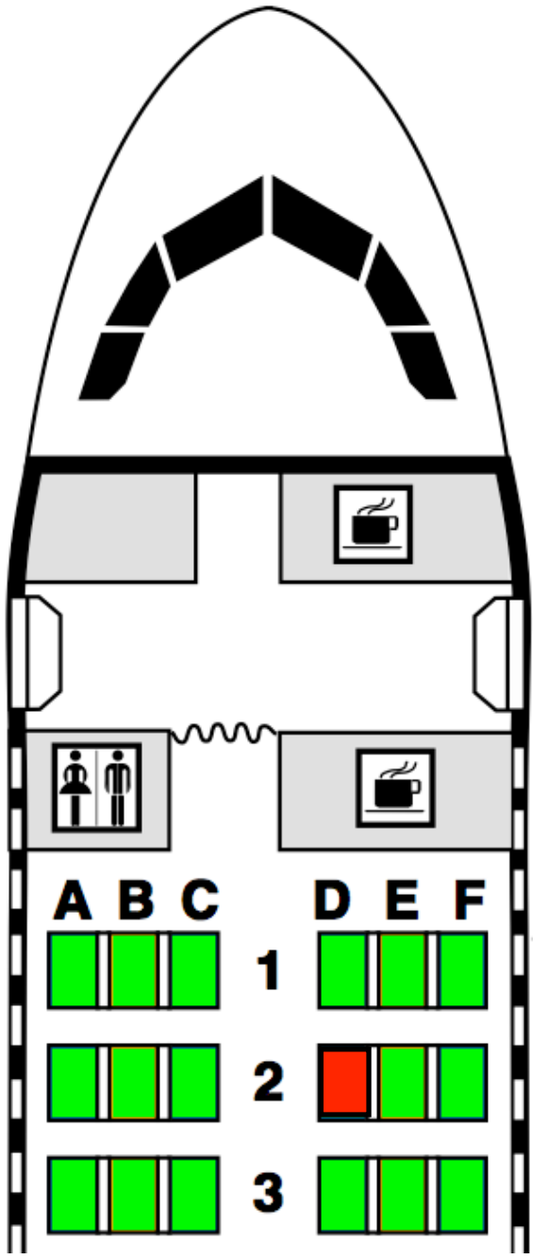
Flugbörse

Reisebüro Meyer

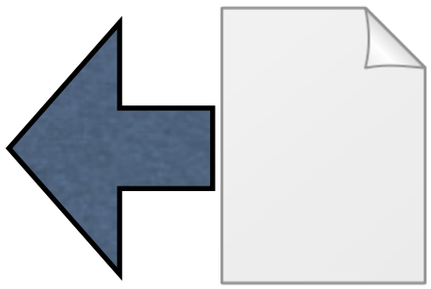
Platz 2D: Alexandra Albrecht

Platz 2D: Andreas Heß

Zugriffskonflikt



Reservierungsdatei zur Bearbeitung sperren (?)



Reservierungsdatei

Flugbörse

Reisebüro Meyer

Platz 2D: Alexandra Albrecht

Platz 2D: Andreas Heß

Beobachtungen (3)

- Weitere erkannte Nachteile:
  - Probleme beim gleichzeitigen Zugriff
  - Konsistenz/Plausibilität nicht sichergestellt
  - Datenverlust droht: *Lost Update*

Werden diese Probleme durch den Einsatz von Datenbanken komplett gelöst?

Motivation
Zusammenfassung

- Vermeidung von
  - unkontrollierter Redundanz
  - Problemen im Mehrbenutzer-Betrieb
    - Datenverlust
    - Inkonsistenzen
- Bessere Zugriffsmöglichkeiten, Verknüpfung von Daten, Rechteverwaltung

vgl. Kemper, S. 19-21

Motivation
Zusammenfassung

- Vermeidung von
  - unkontrollierter Red
  - Problemen im Mehrb
  - Datenverlust
  - Inkonsistenzen

Ohne Datenbanken: Programmierer selbst verantwortlich für Einhaltung der Konsistenz sowie Erfüllung aller weiteren Anforderungen

- Bessere Zugriffsmöglichkeiten, Verknüpfung von Daten, Rechteverwaltung

vgl. Kemper, S. 19-21

Motivation
Zusammenfassung

- Vermeidung von
  - unkontrollierter Red
  - Problemen im Mehrb
  - Datenverlust
  - Inkonsistenzen
- Bessere Zugriffsmöglichkeiten, Verknüpfung von Daten, Rechteverwaltung
- Geringere Entwicklungskosten

Ohne Datenbanken: Programmierer selbst verantwortlich für Einhaltung der Konsistenz sowie Erfüllung aller weiteren Anforderungen

vgl. Kemper, S. 19-21

Anforderungen an Datenbanken (1)

- Grundlegende Funktionen:
  - Lesen, Schreiben, Ändern von Daten
  - Abstraktion der physischen Datenzugriffe

Abstraktion?
Was bedeutet das, was bringt das?

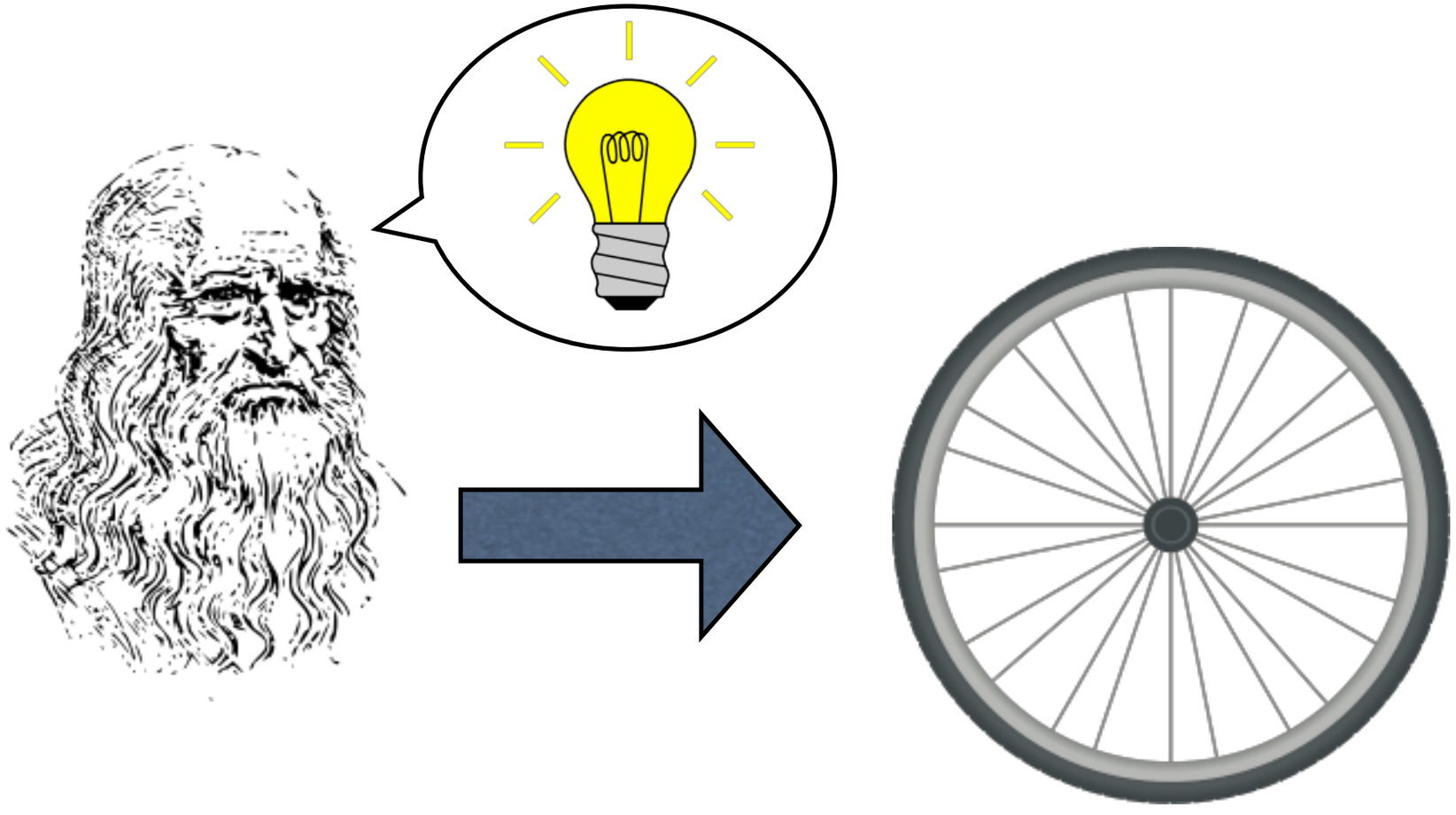
Abstraktion

Abstraktion

- Zugriff auf Zettelkasten nur durch Bibliothekarin!
- Vereinfachung für den Nutzer

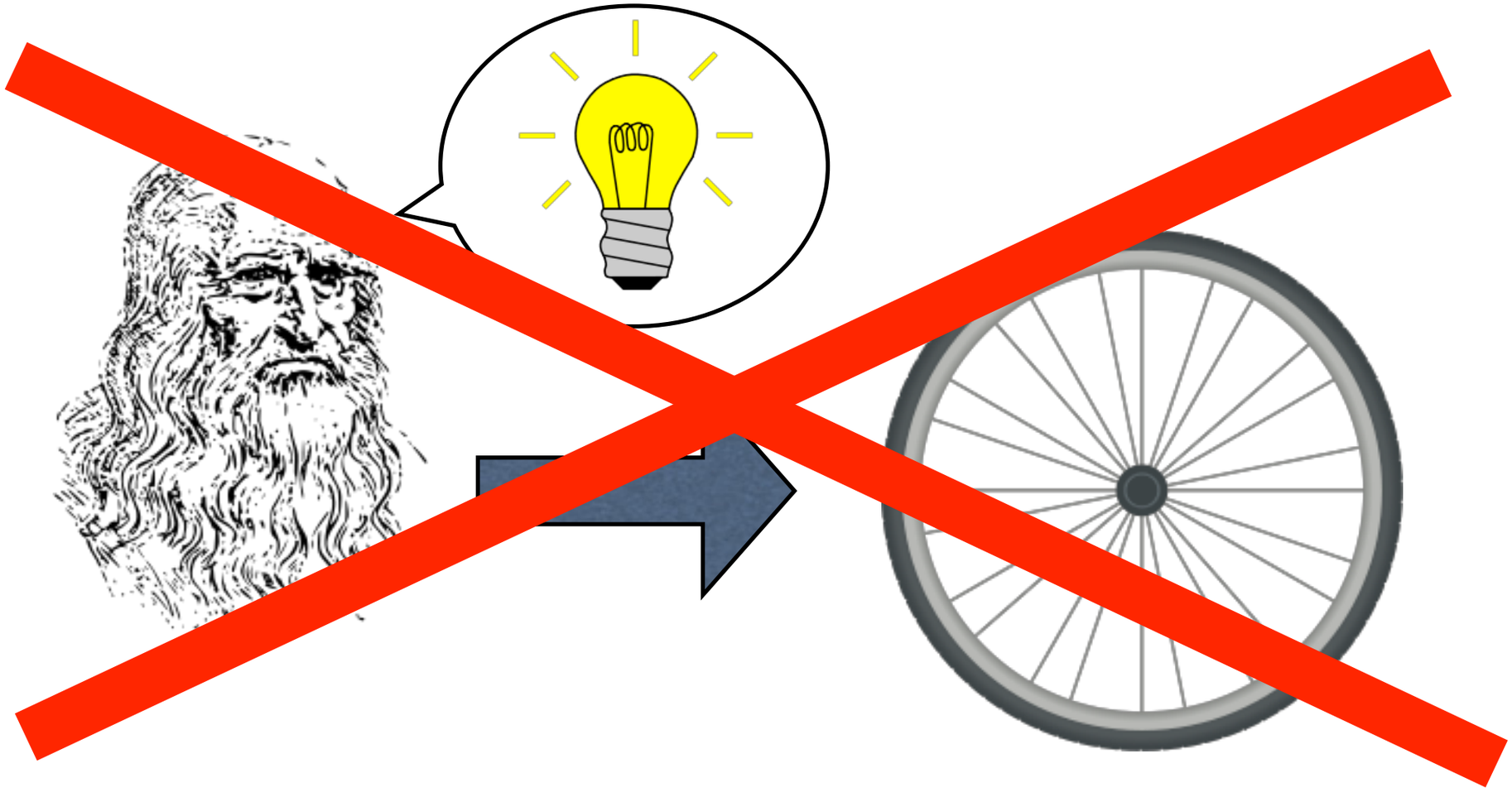
Abstraktion

- Programmierer soll sich nicht jedes Mal aufs Neue mit Grundfunktionen befassen



Abstraktion

- Programmierer soll sich nicht jedes Mal aufs Neue mit Grundfunktionen befassen



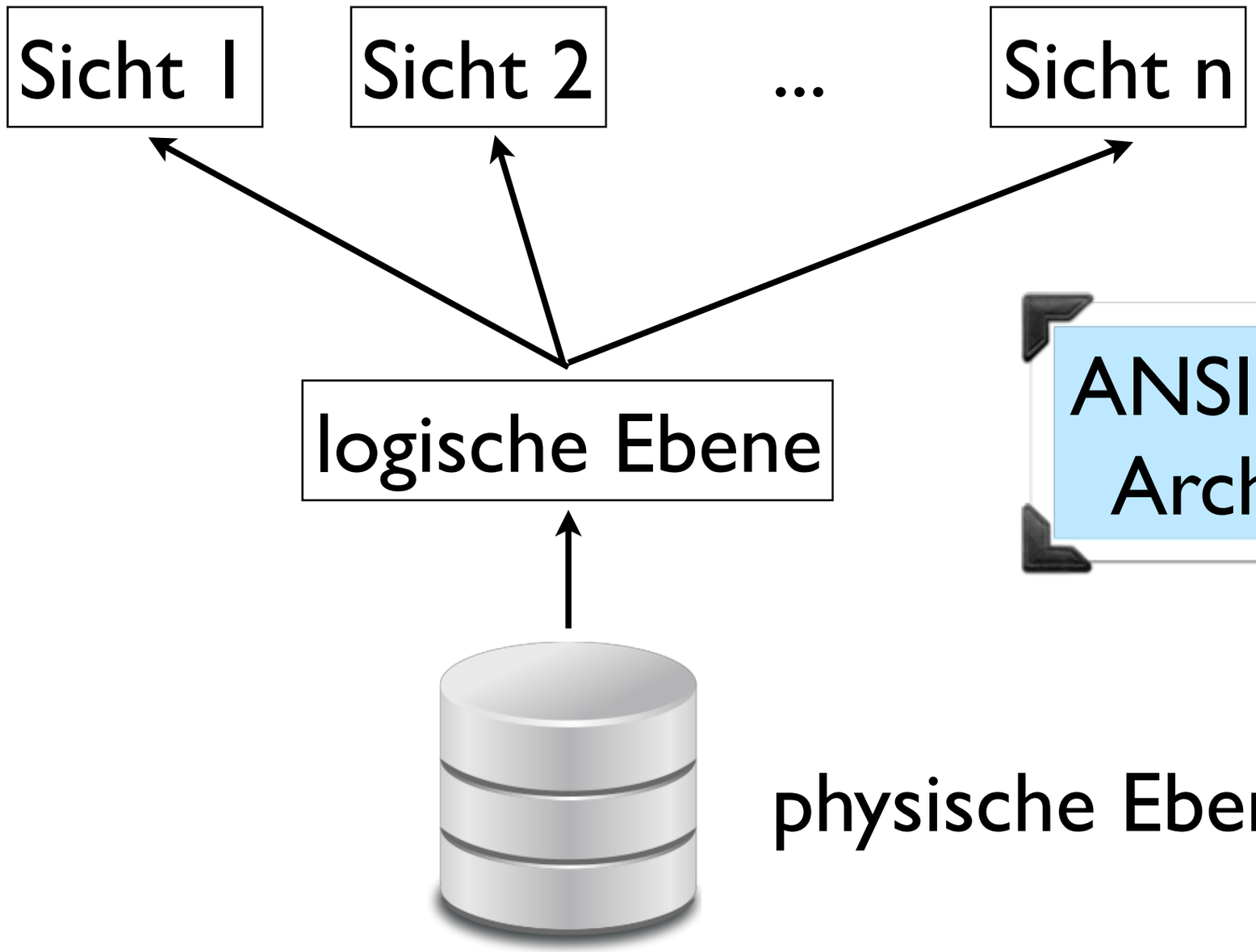
Abstraktion

- Programmierer soll sich nicht jedes Mal aufs Neue mit Grundfunktionen befassen

Die Gründe:

- Wer programmiert, macht Fehler!
- Wiederverwertung spart Zeit & Kosten

Abstraktion



ANSI-SPARC-Architektur

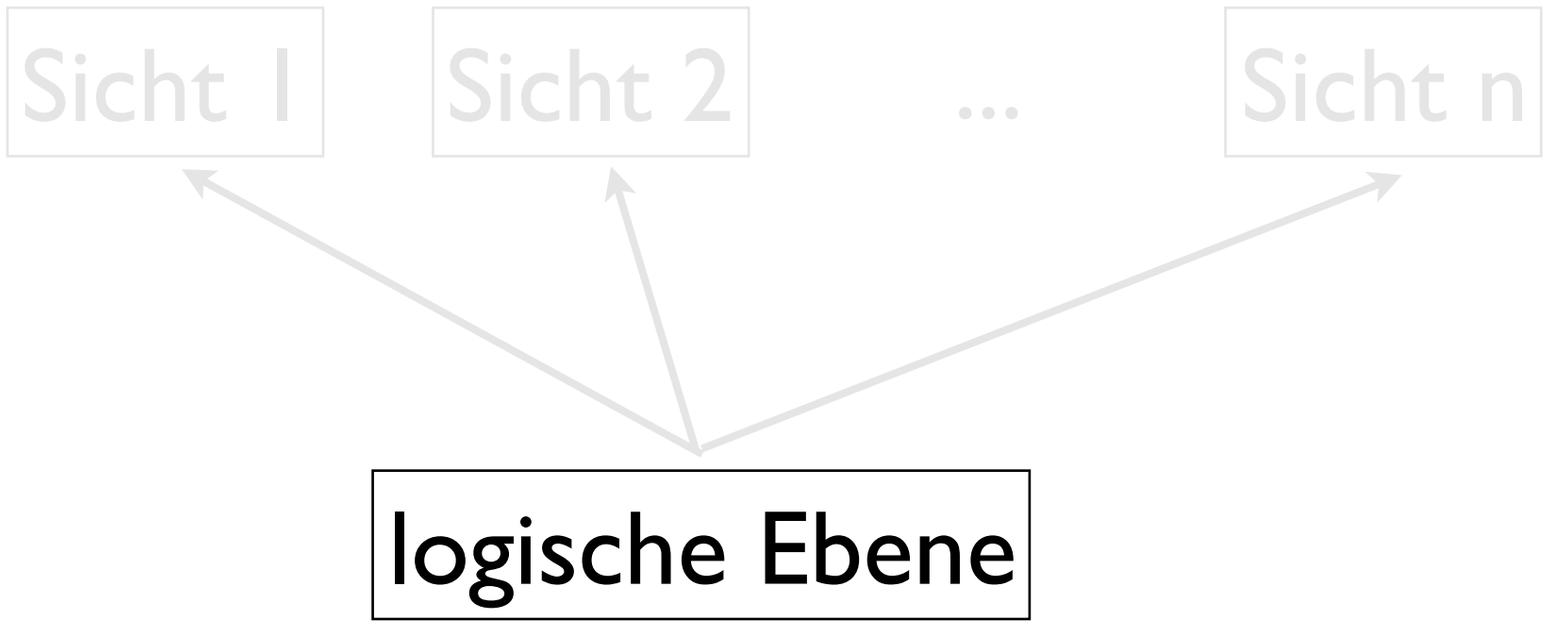
vgl. Kemper, S. 21-22

Physische Ebene

- Wie sind die Daten gespeichert
  - Speicherstrukturen usw.
  - nicht für den „normalen Benutzer“
  - was konkret auf der Festplatte liegt

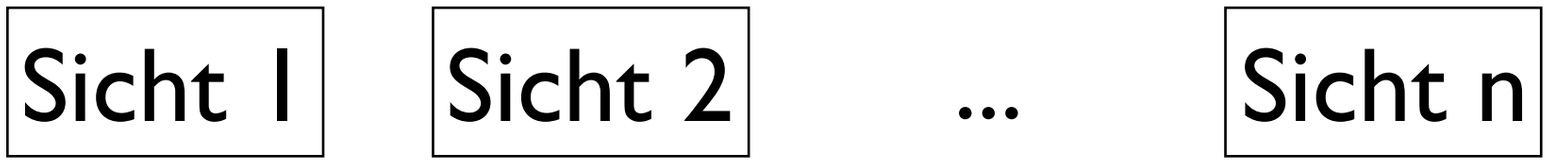
physische Ebene

Logische Ebene



- Welche Daten können gespeichert werden
  - *Datenbankschema*
    legt die Struktur der Daten fest

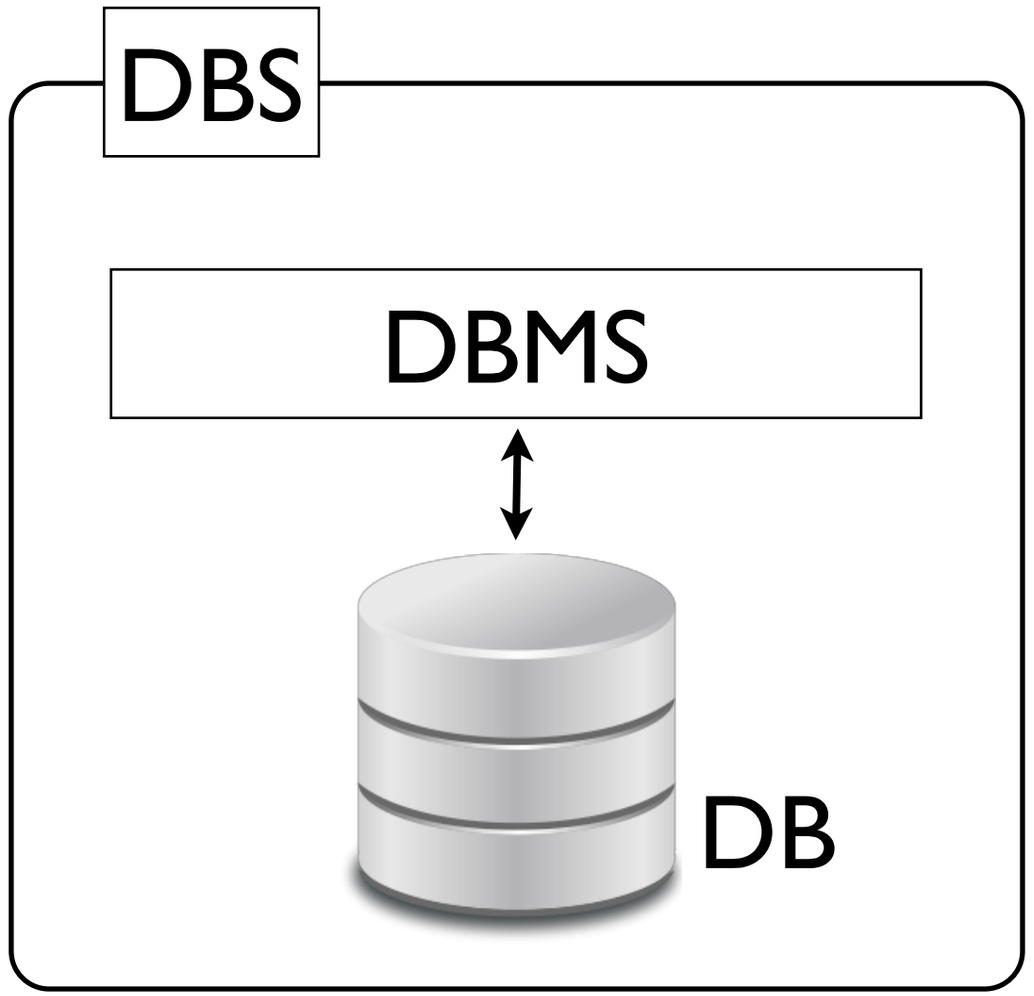
Sichten



- Was sieht das Anwendungsprogramm
- *Sichten* stellen Teilmenge der Informationen bereit
- Zugeschnitten auf Bedürfnisse der Anwendung
- Meistens nur zum Lesen

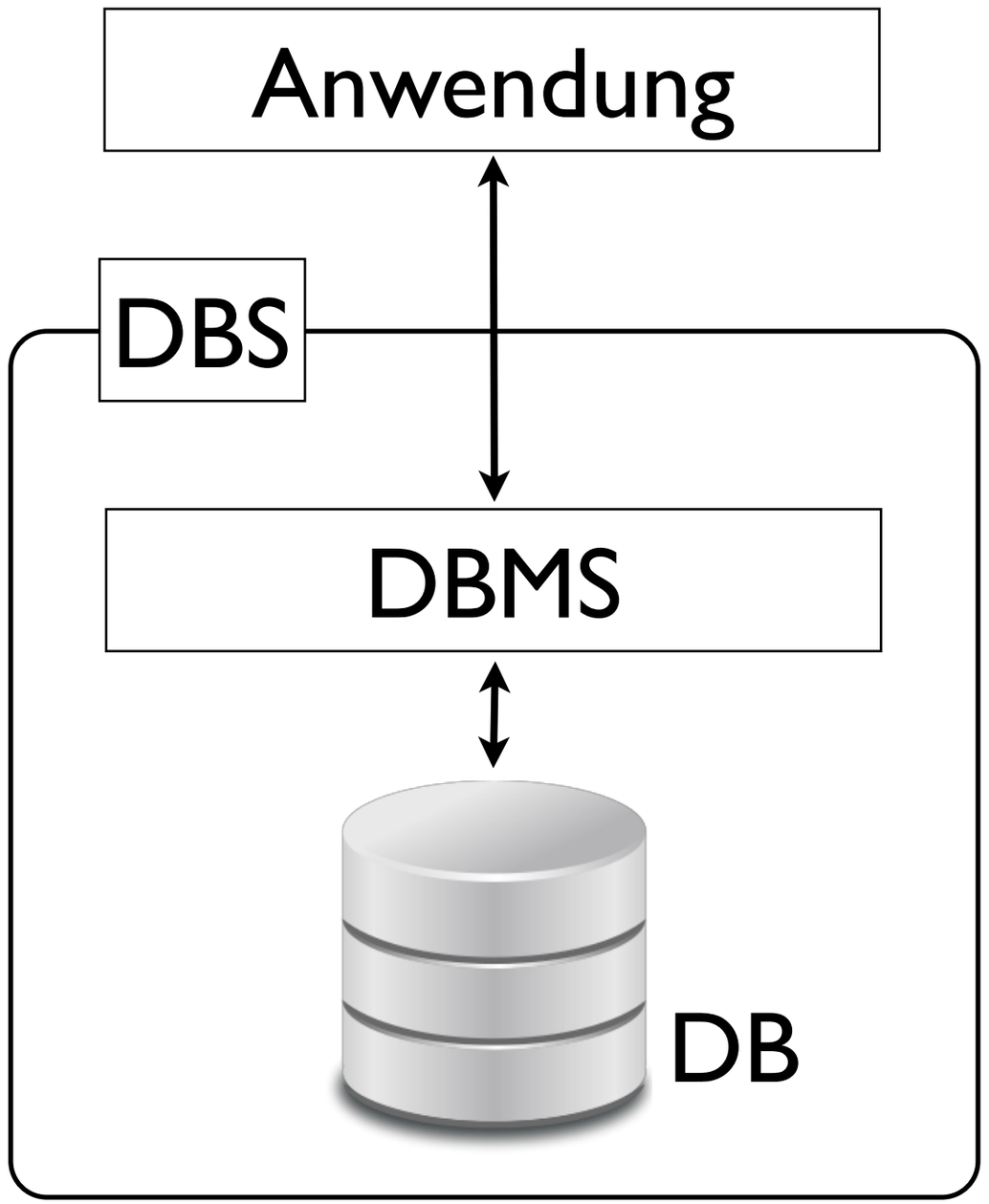
Begriffsklärung (I)

- *Datenbank-Management-System (DBMS):*
  Software zur Verwaltung der Daten
- *Datenbank (DB):*
  Ansammlung von Daten, vom DBMS verwaltet
- *Datenbank-System (DBS):*
  DBMS + DB



Begriffsklärung (I)

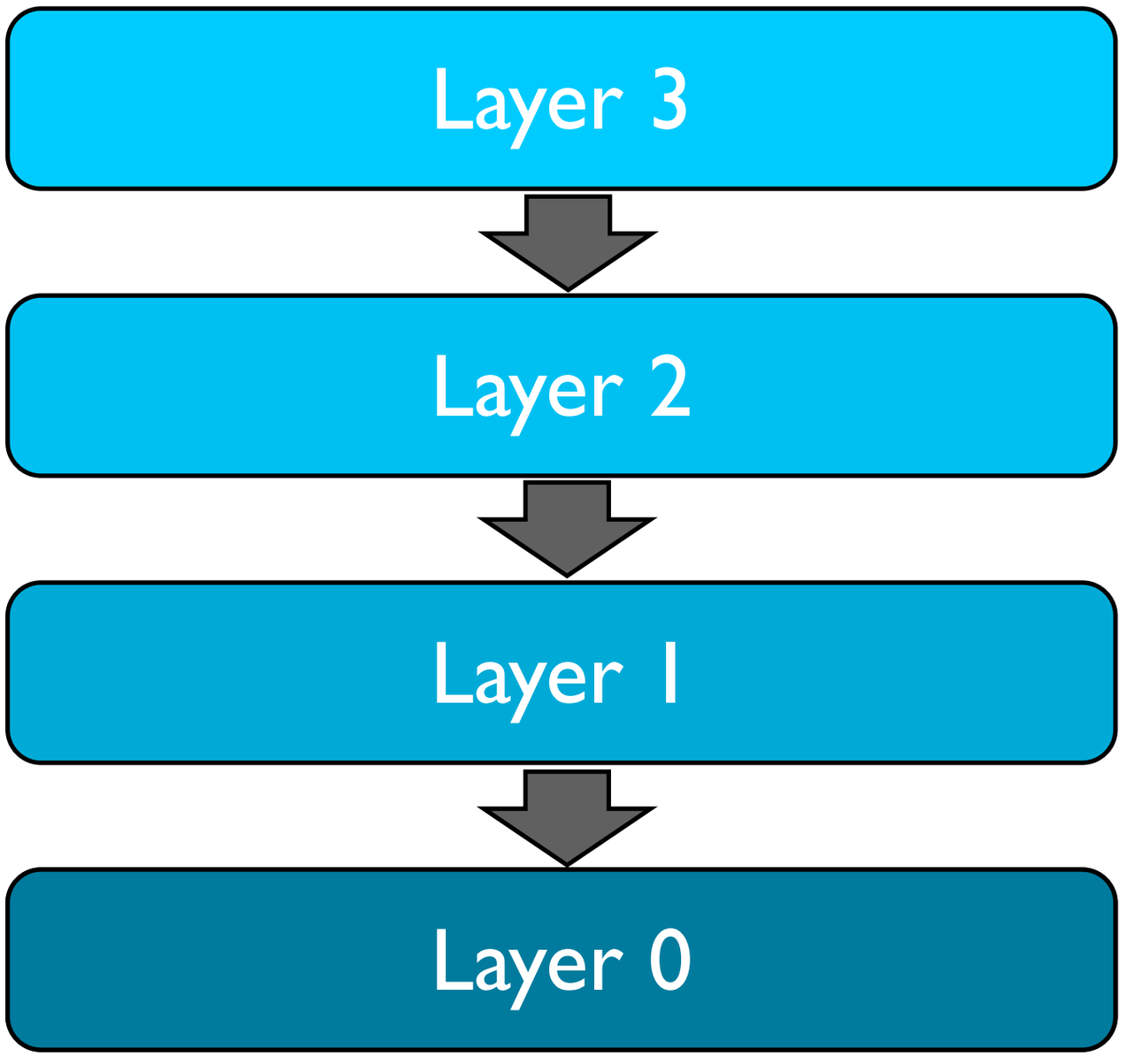
- *Datenbank-Management-System (DBMS):*
  Software zur Verwaltung der Daten
- *Datenbank (DB):*
  Ansammlung von Daten, vom DBMS verwaltet
- *Datenbank-System (DBS):*
  DBMS + DB



Begriffsklärung (2)

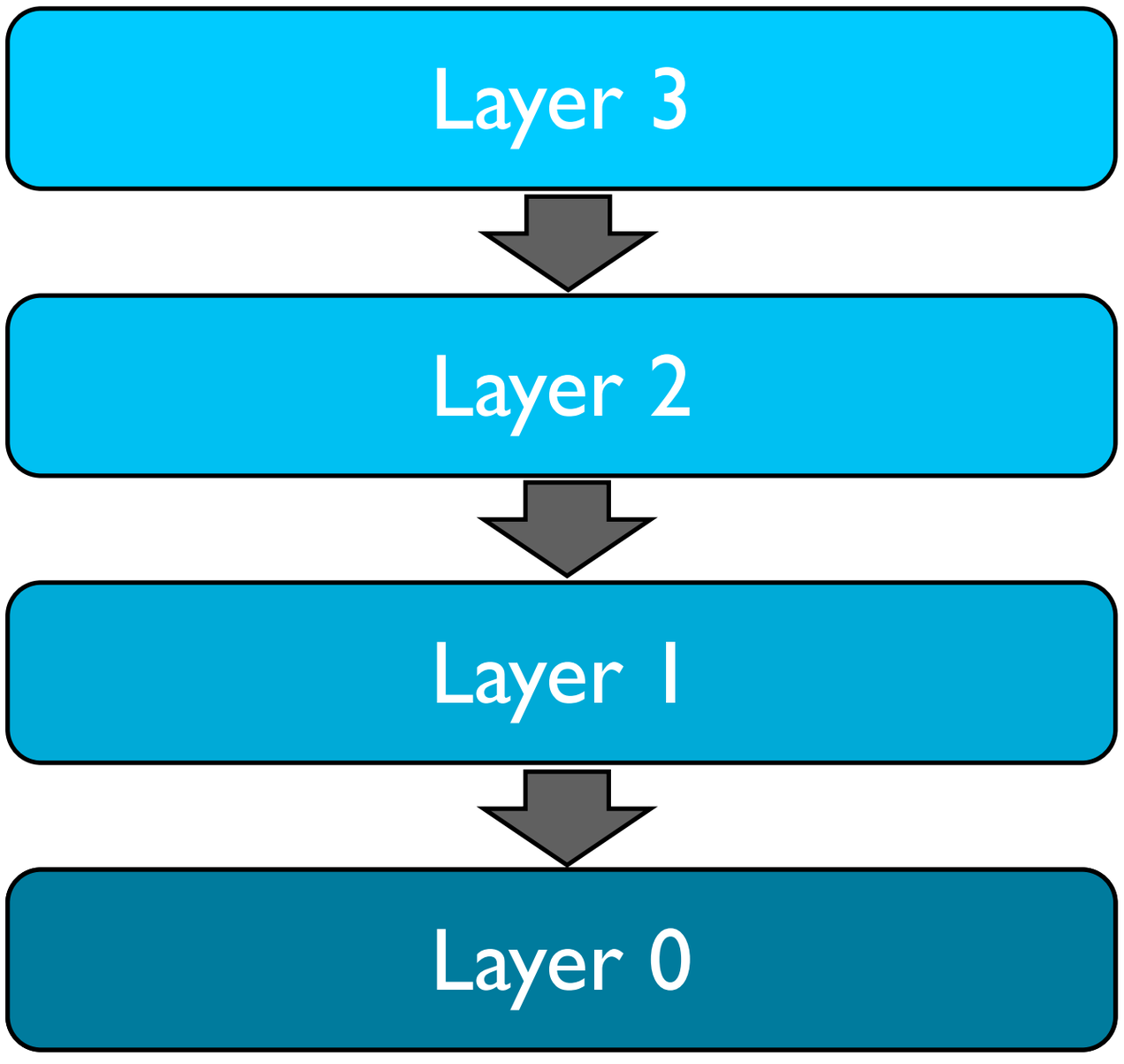
- *Datenunabhängigkeit:*
  - *physische Datenunabhängigkeit:*
    Änderungen auf der physischen Ebene berührt die logische Ebene nicht
  - *logische Datenunabhängigkeit:*
    Änderungen auf der logischen Ebene berühren die Anwendungen nicht

Exkurs: Schichtenmodelle



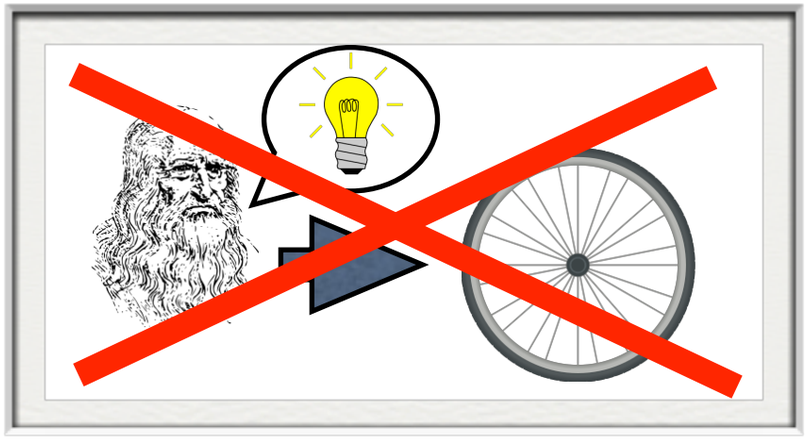
- Jede Schicht: Programm-Teile
- Kann darunter liegende Schichten benutzen

Exkurs: Schichtenmodelle

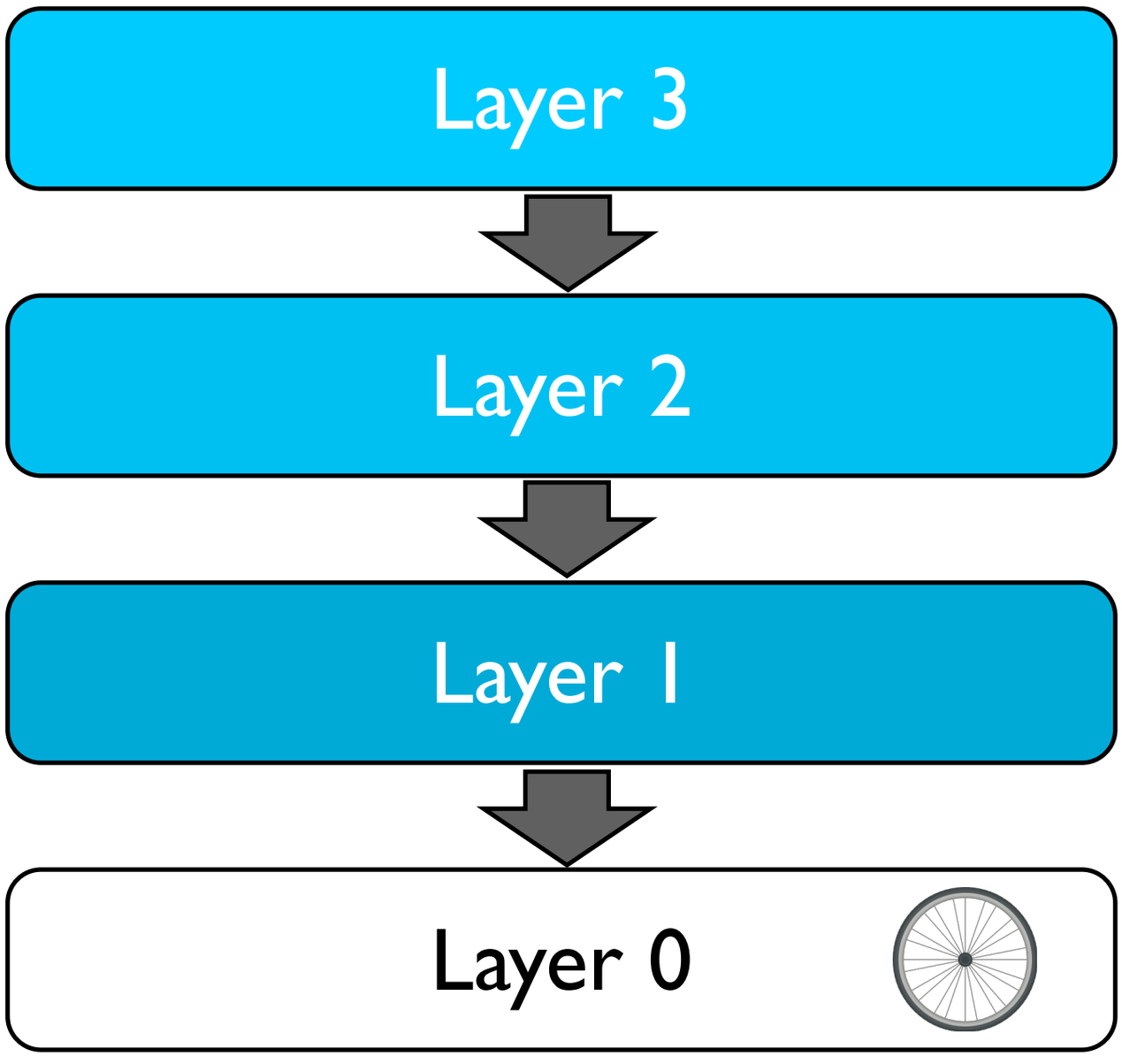


- Jede Schicht: Programm-Teile
- Kann darunter liegende Schichten benutzen

Der Grund:

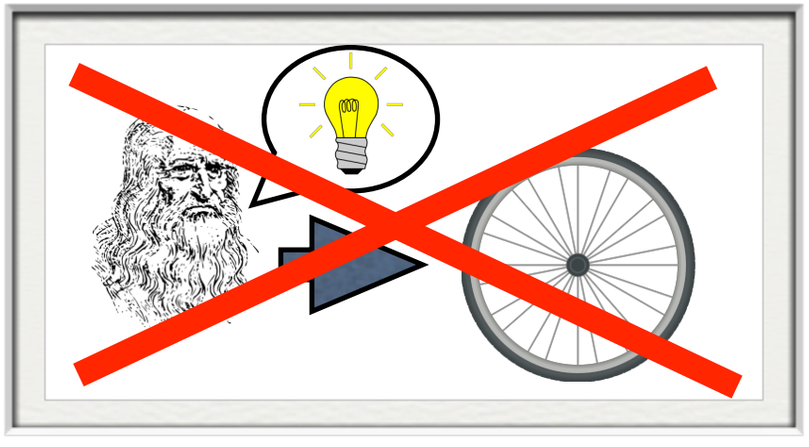


Exkurs: Schichtenmodelle

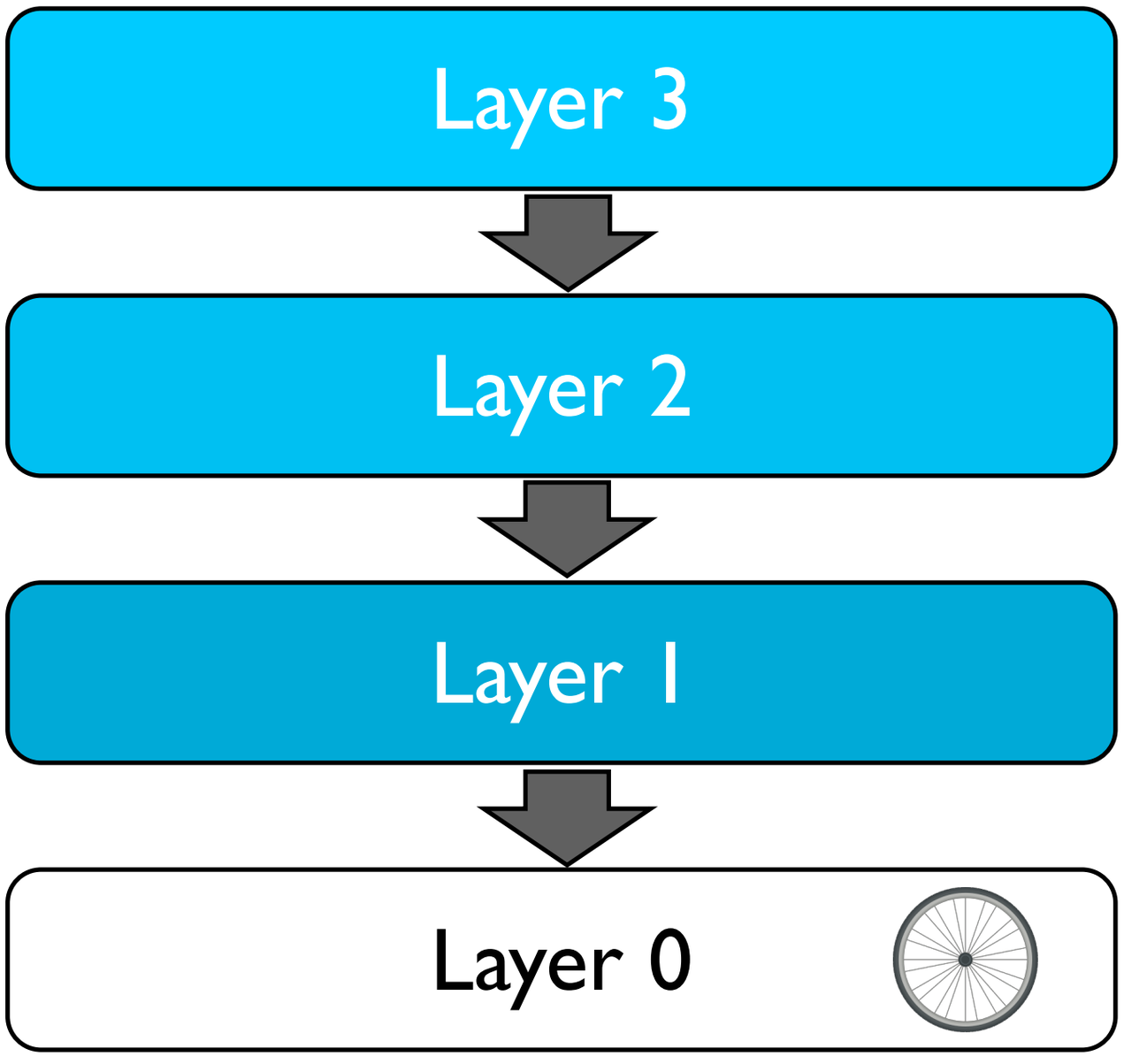


- Jede Schicht: Programm-Teile
- Kann darunter liegende Schichten benutzen

Der Grund:

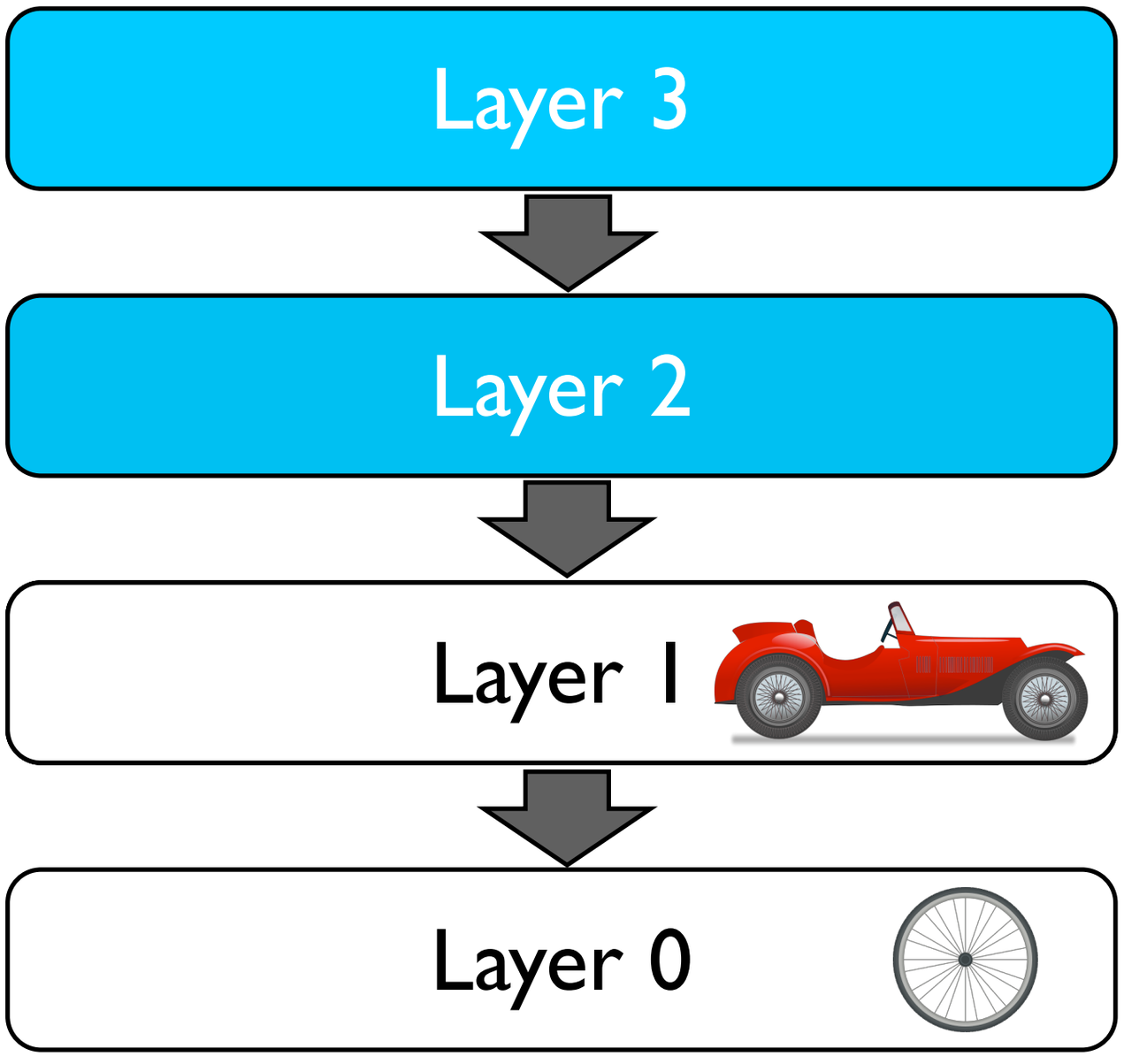


Exkurs: Schichtenmodelle



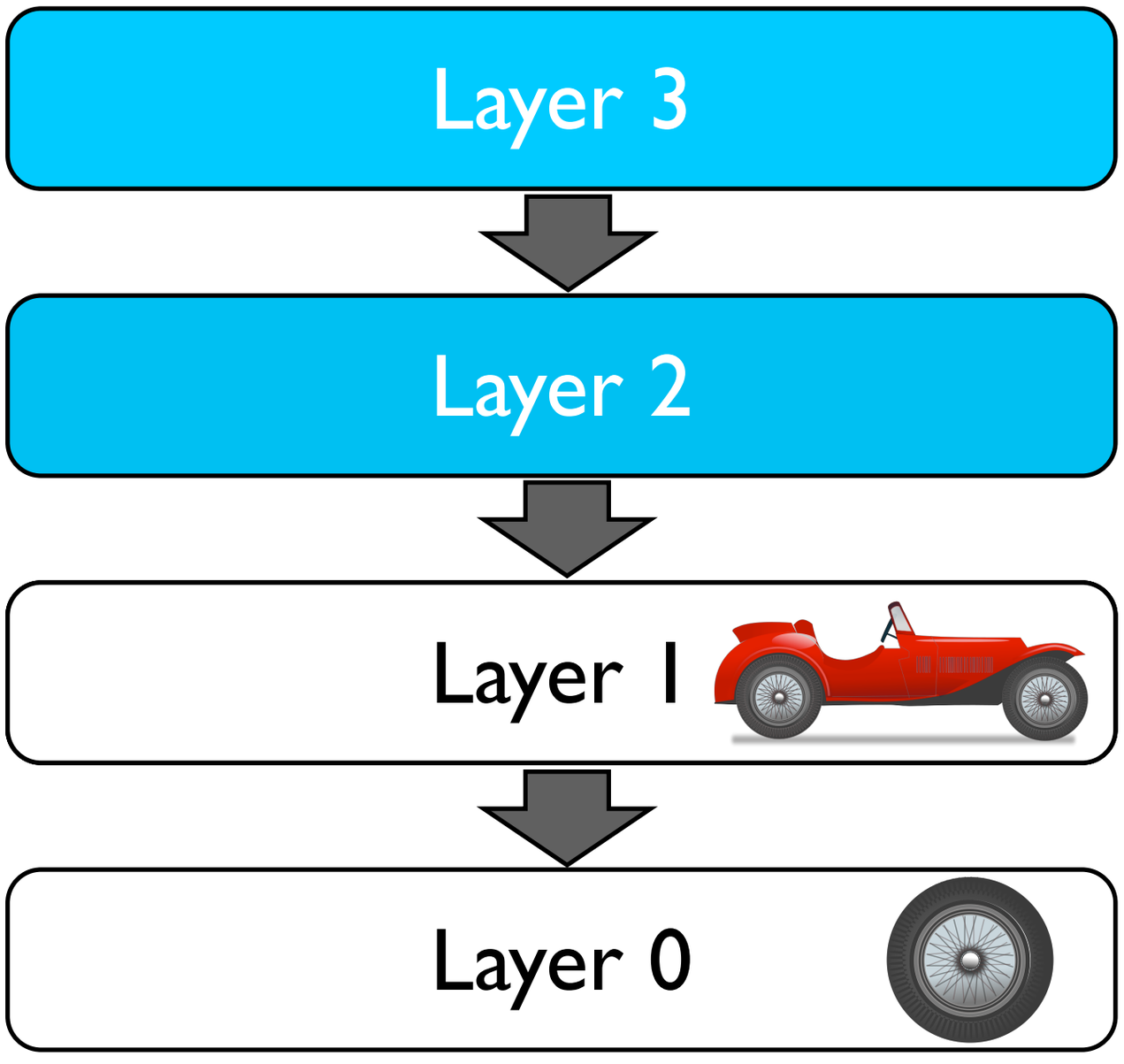
- Was passiert mit dem Programm in Schicht 1, wenn sich das Programm in Schicht 0 ändert?

Exkurs: Schichtenmodelle



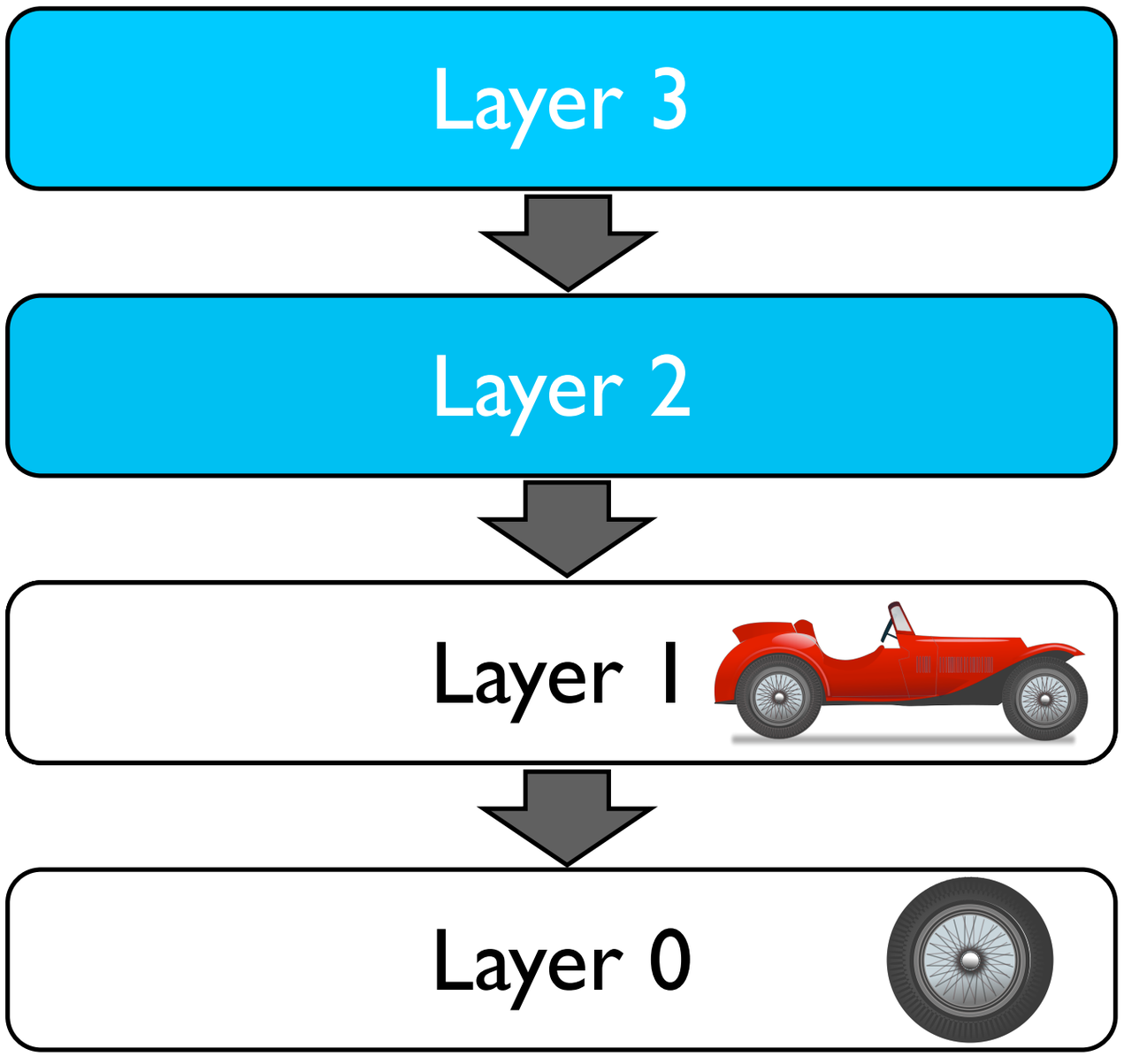
- Was passiert mit dem Programm in Schicht 1, wenn sich das Programm in Schicht 0 ändert?

Exkurs: Schichtenmodelle



- Was passiert mit dem Programm in Schicht 1, wenn sich das Programm in Schicht 0 ändert?

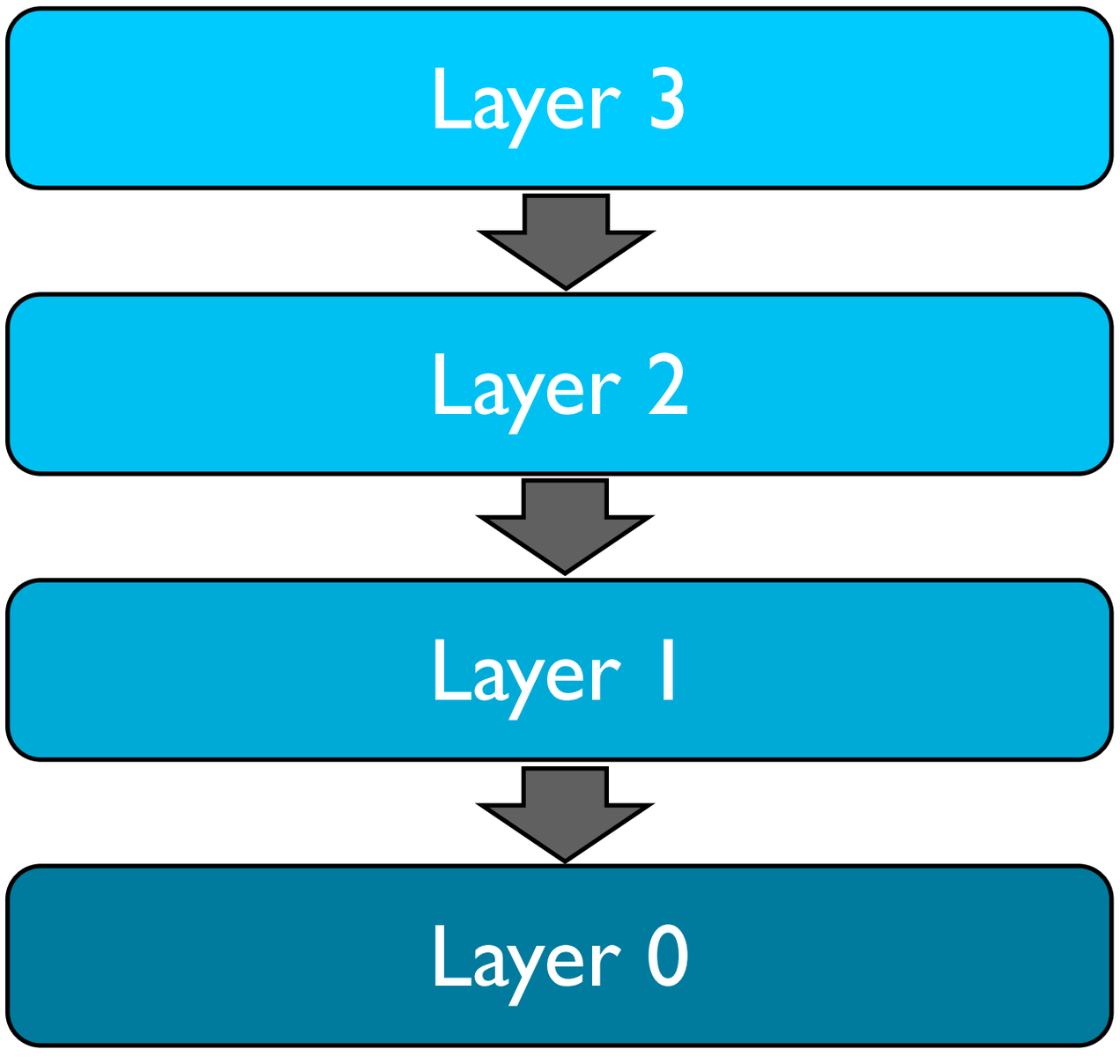
Exkurs: Schichtenmodelle



- Was passiert mit dem Programm in Schicht 1, wenn sich das Programm in Schicht 0 ändert?

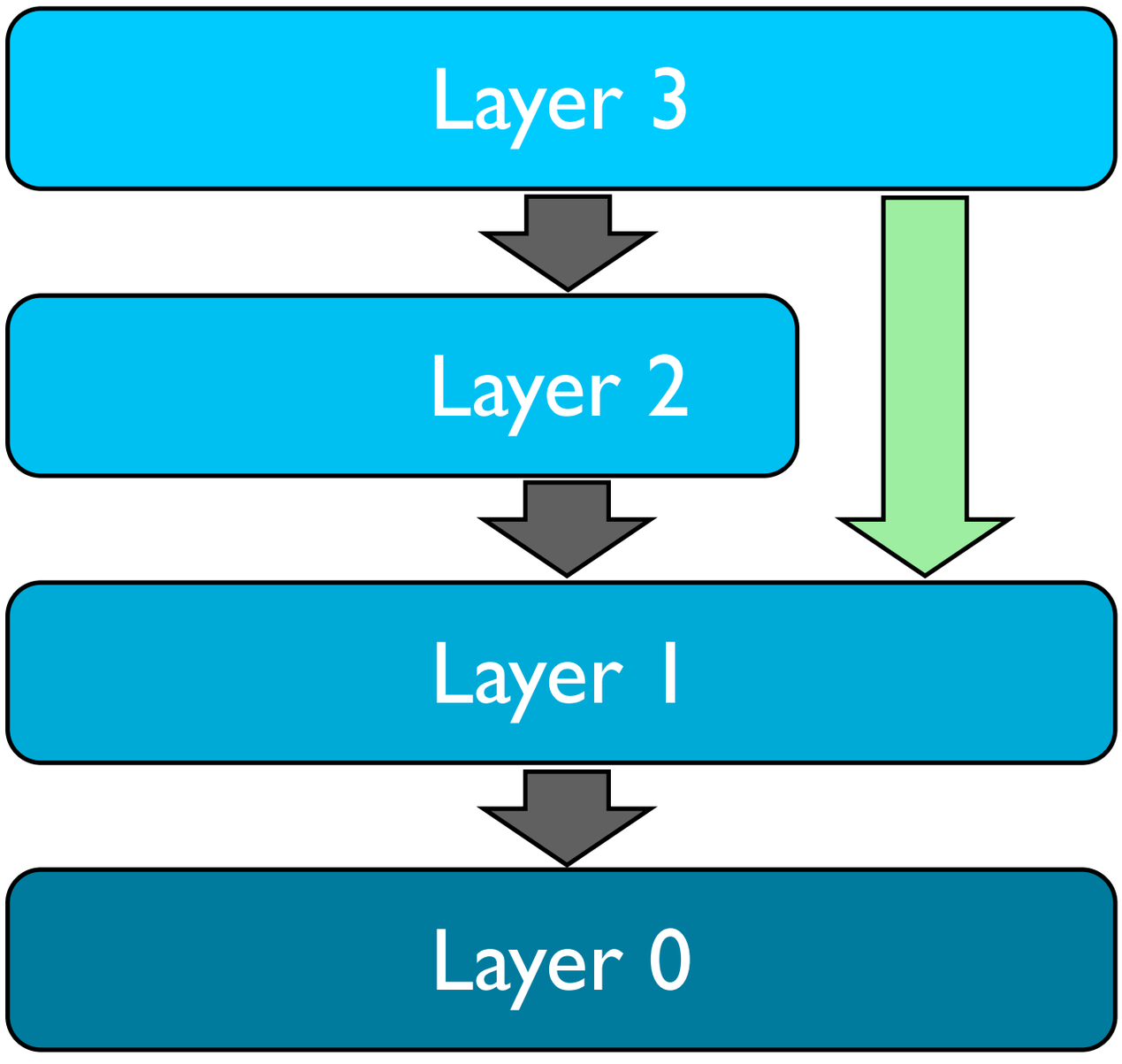
Idealerweise ändert sich nichts!

Exkurs: Schichtenmodelle

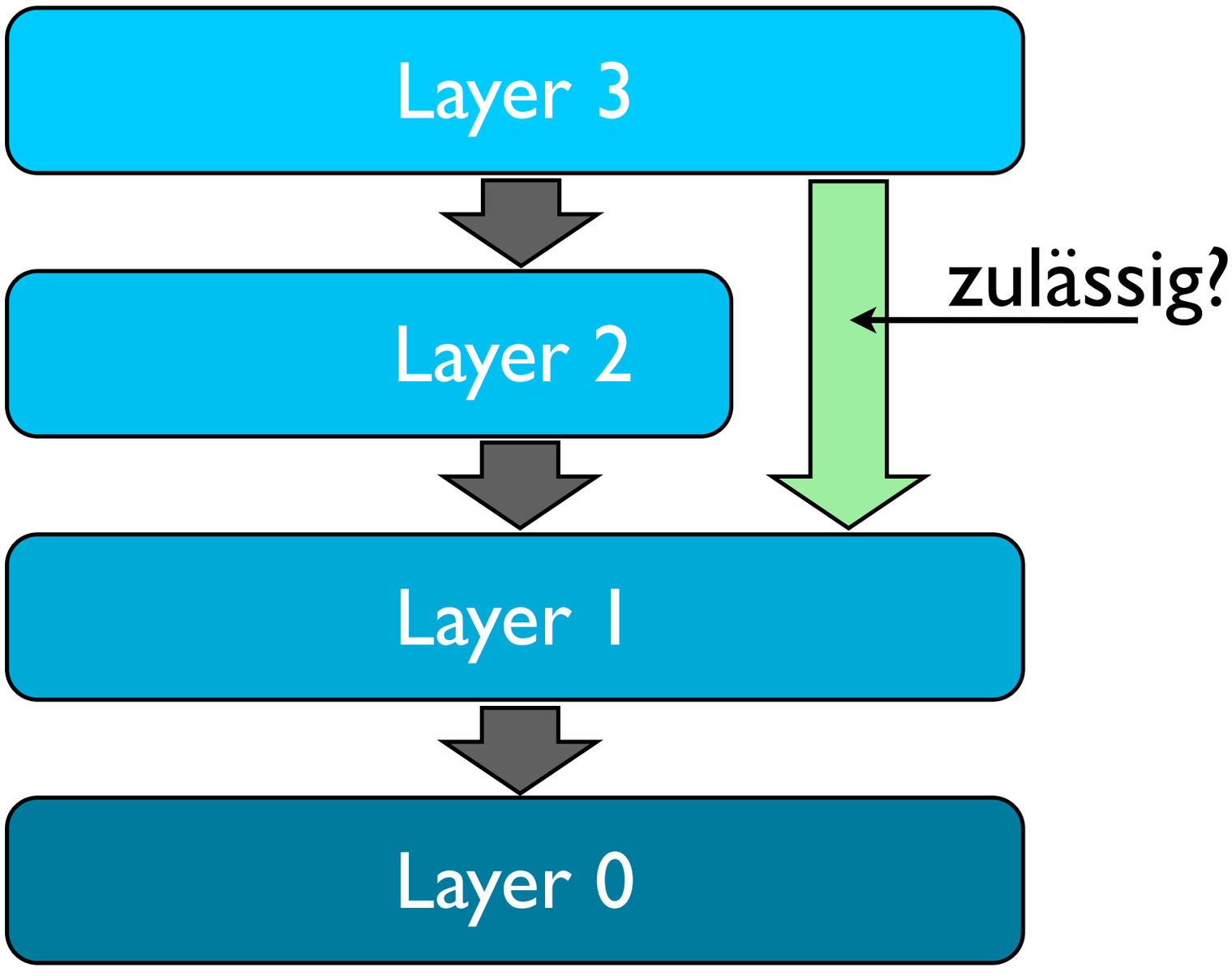


- Jede Schicht sieht nur <u>darunter</u> liegende Schichten
- Ziele:
  - Kapselung von Funktionen
  - Verbergen von Interna

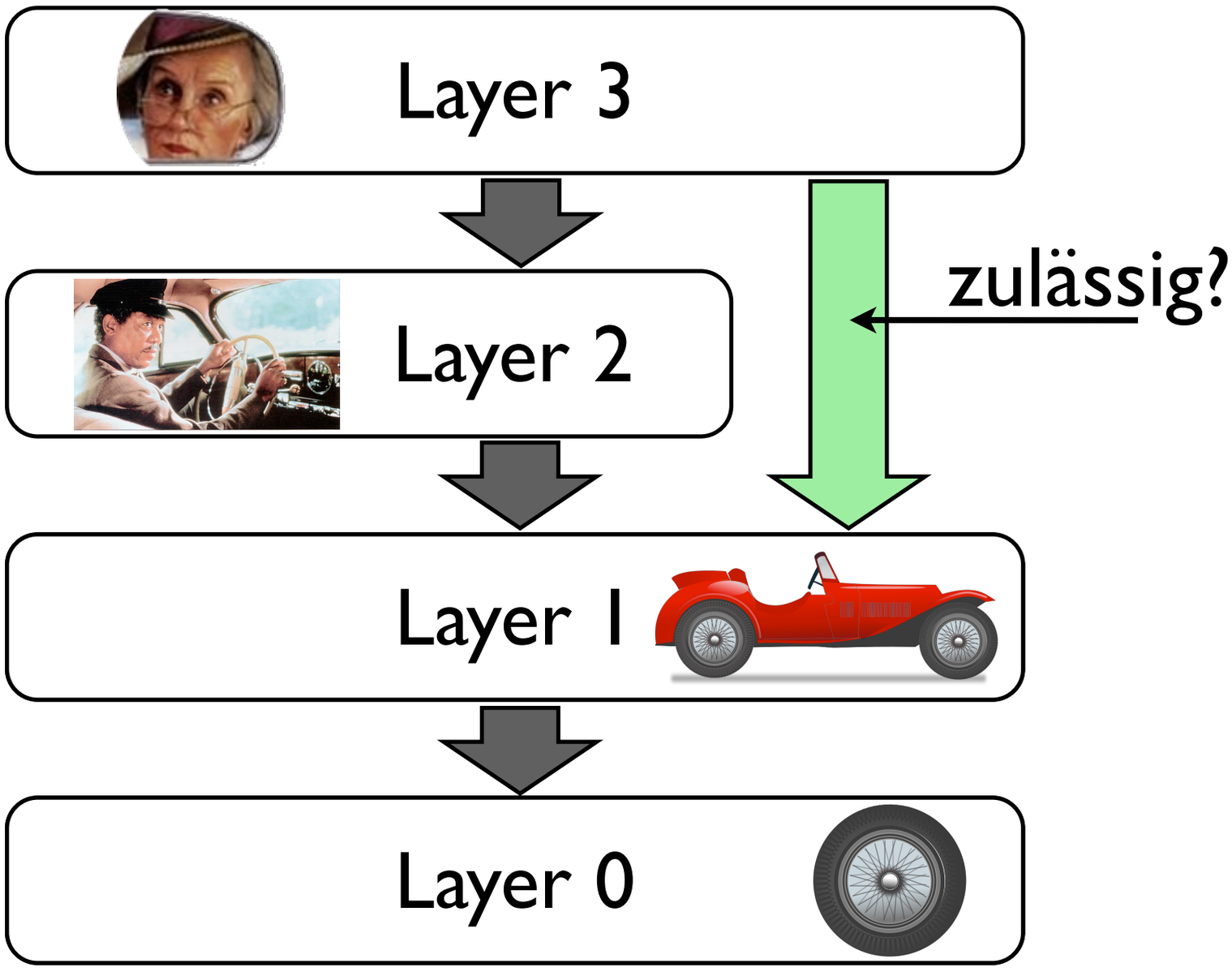
Exkurs: Schichtenmodelle



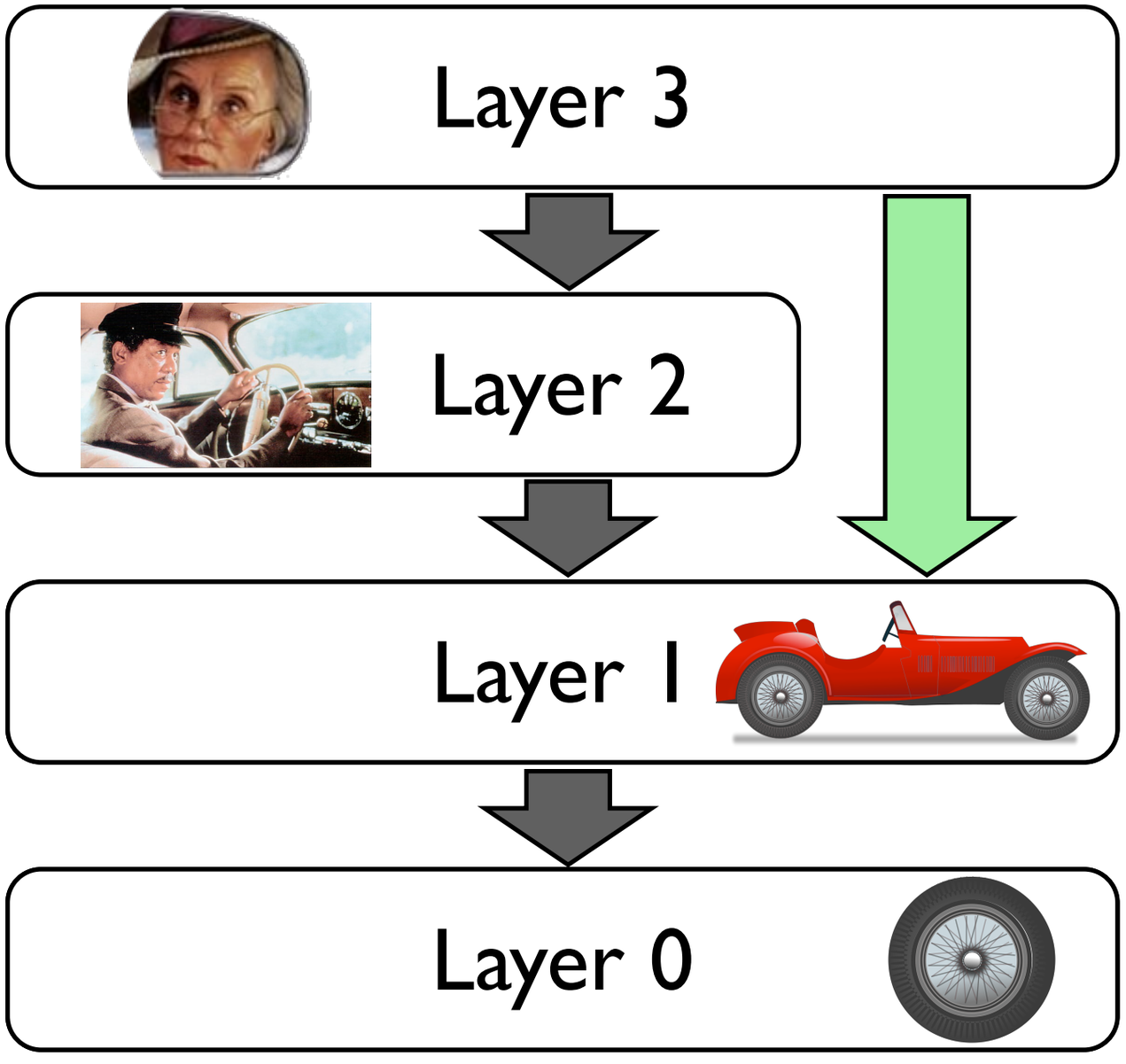
Exkurs: Schichtenmodelle



Exkurs: Schichtenmodelle

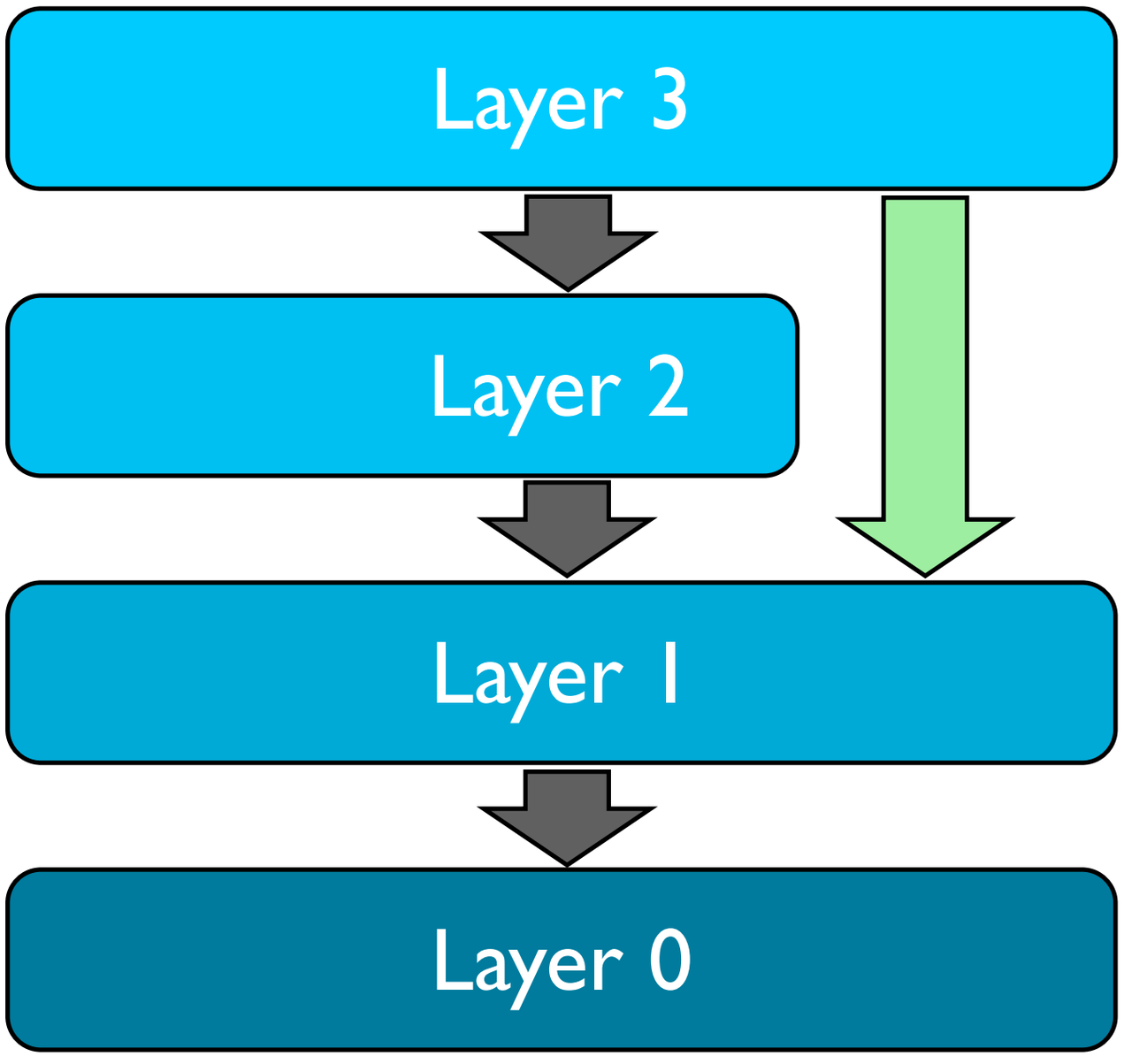


Exkurs: Schichtenmodelle



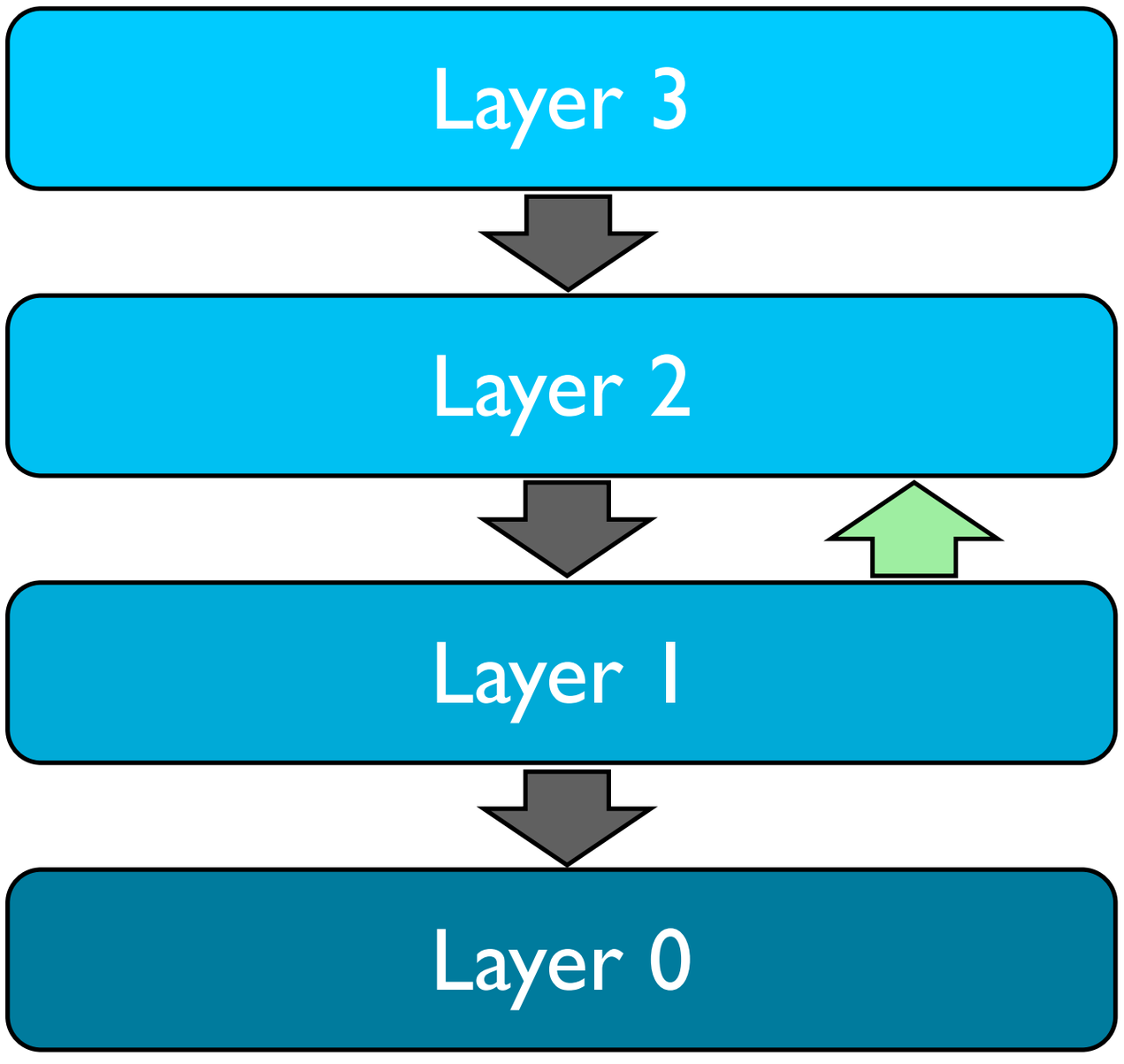
- Bei striktem Schichtenmodell nur direkt angrenzende Schichten, sonst ok

Exkurs: Schichtenmodelle

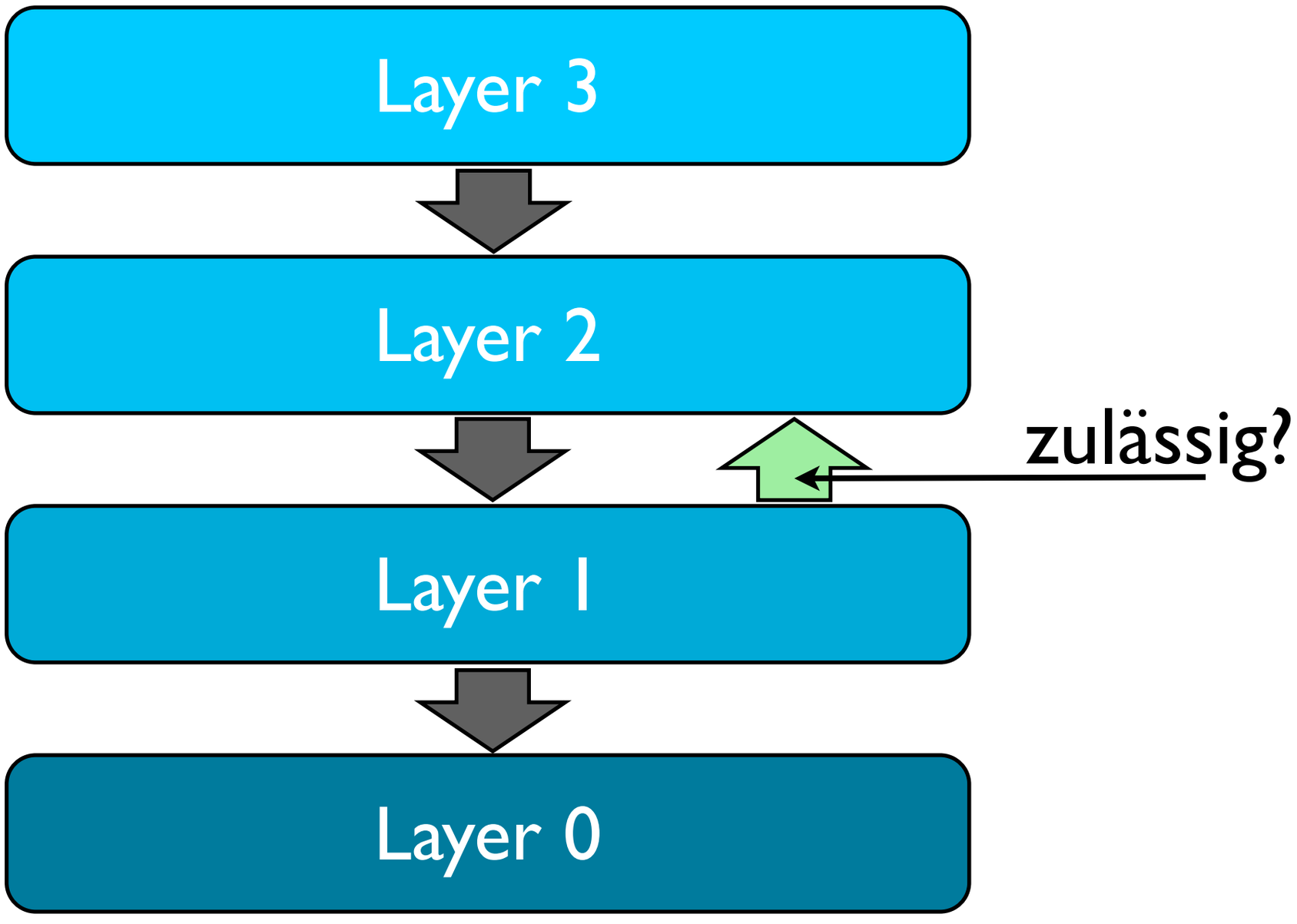


- Bei striktem Schichtenmodell nur direkt angrenzende Schichten, sonst ok

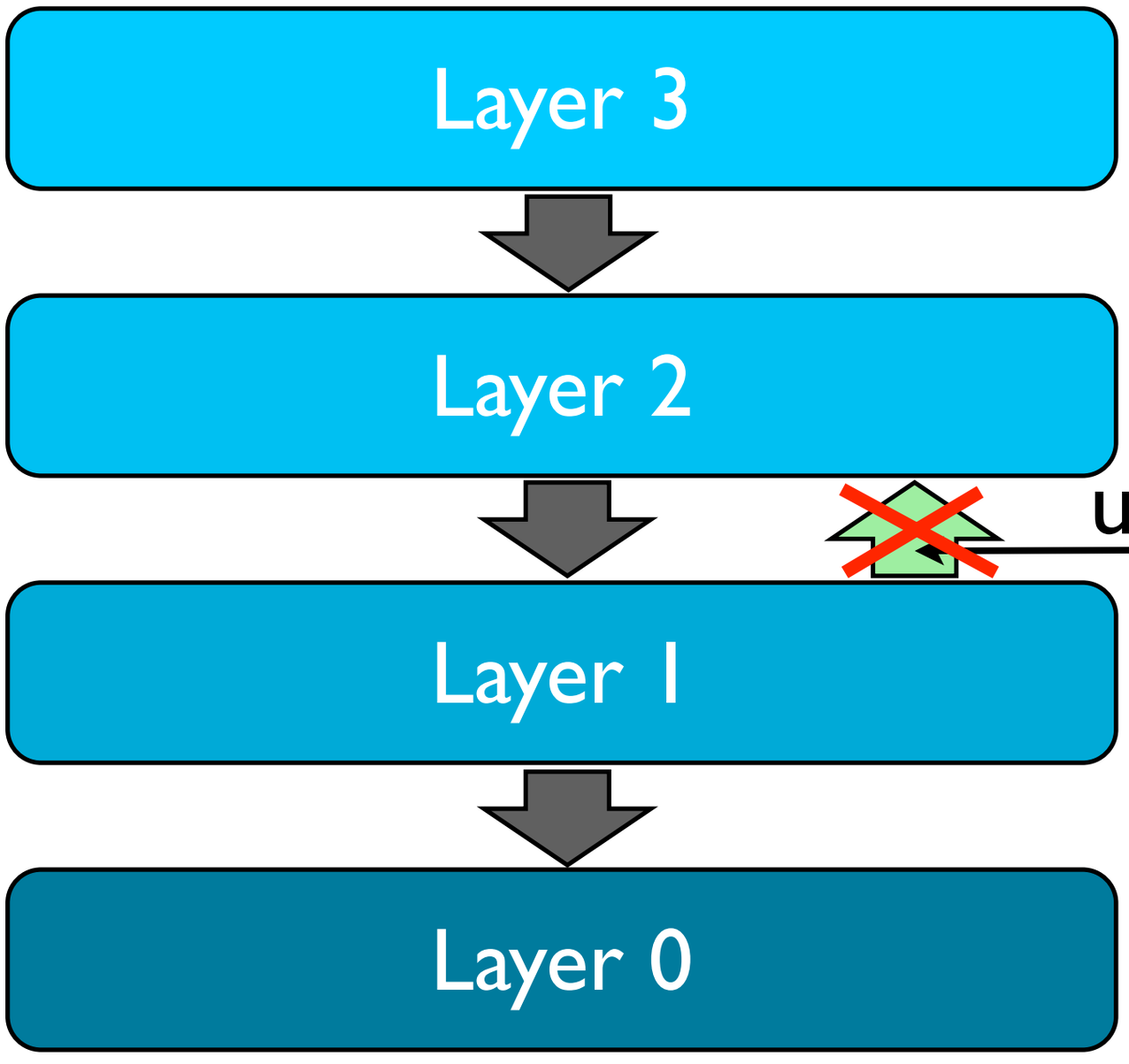
Exkurs: Schichtenmodelle



Exkurs: Schichtenmodelle



Exkurs: Schichtenmodelle



- Jede Schicht sieht nur <u>darunter</u> liegende Schichten

Anforderungen an DBMS (2)

kurze Rückblende...

Motivation
Zusammenfassung

- Vermeidung von
  - unkontrollierter Redundanz
  - Problemen im Mehrbenutzer-Betrieb
    - Datenverlust
    - Inkonsistenzen
- Bessere Zugriffsmöglichkeiten, Verknüpfung von Daten, Rechteverwaltung
- Geringere Entwicklungskosten

Anforderungen an DBMS (2)

- DBMS soll unterstützen bei
  - Integrität, Konsistenz sichern
  - Mehrbenutzerbetrieb
  - Vermeidung von unerwünschter Redundanz
  - Verwaltung von Zugriffsrechten
  - Verknüpfung von Daten

Was braucht ein DBMS?

- Datenmodell
- Beschreibung der Datenobjekte
- Anwendbare Operatoren und Wirkung

vgl. Kemper, S. 23

Begriffsklärung (3)

- *Data Definition Language (DDL):*
  Beschreibt Struktur der Daten
- *Data Manipulation Language (DML):*
  besteht aus
  - *Query Language* zum Lesen von Daten
  - Sprache zum Einfügen/Ändern/Löschen

Begriffsklärung (4)

- *Datenbankschema:*
  legt die Struktur der abspeicherbaren Daten fest
- *Datenbankausprägung:*
  die abgespeicherten Daten selbst

vgl. Kemper, S. 24

Entwurf einer DB

- Festlegen eines Datenbankschemas

Kemper	Alfons	Schillerstraße 15	55116	Mainz	kemper@example.com
Meier	Hans	Goethestraße 23a	12345	Mönchengladbach	hans@gmx.cz
Heß	Andreas	Lessingstr. 25	78120	Furtwangen	hea@hs-furtwangen.de
Albrecht	Alexandra	Marktplatz 1	67543	Heidelberg	alex@albrecht.com

Entwurf einer DB

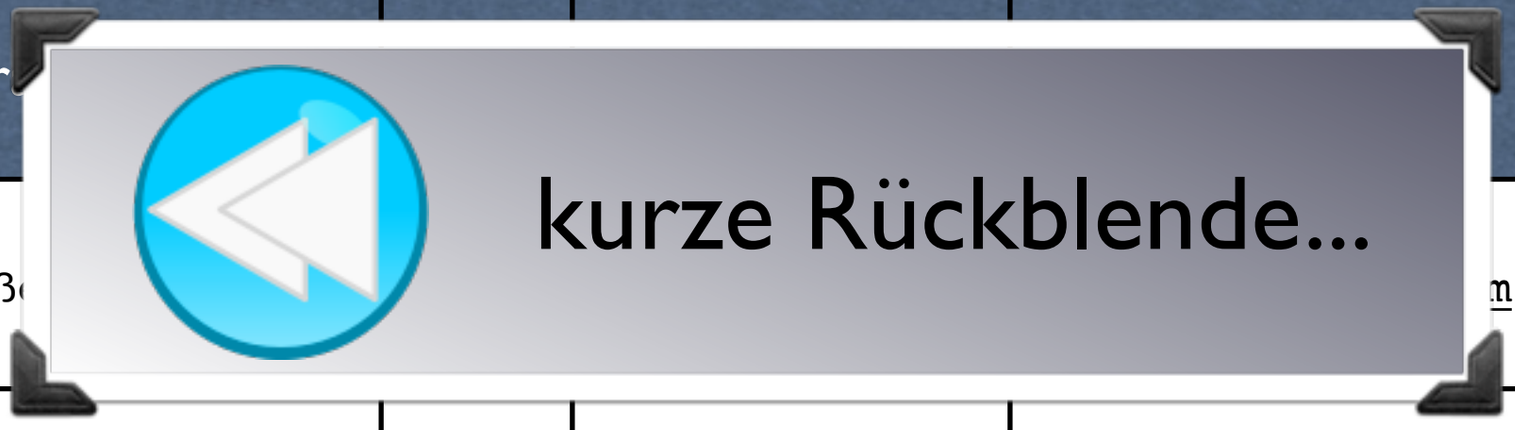
- Wie sehen die Daten in den Spalten aus?

Name	Vorname	Straße	PLZ	Ort	E-Mail
Kemper	Alfons	Schillerstraße 15	55116	Mainz	kemper@example.com
Meier	Hans	Goethestraße 23a	12345	Mönchengladbach	hans@gmx.cz
Heß	Andreas	Lessingstr. 25	78120	Furtwangen	hea@hs-furtwangen.de
Albrecht	Alexandra	Marktplatz 1	67543	Heidelberg	alex@albrecht.com

Entwurf einer DB

- Wie sehen die Daten in den Spalten aus?

Name	Vorname	Str			
Kemper	Alfons	Schillerstraß			m
M				Mönchengladbach	hans@gmx.cz
H				Furtwangen	hea@hs-furtwangen.de
Albrecht	Alexandra	Marktplatz 1	ABC0	Heidelberg	alex@albrecht.com



„Aber in eine Datei kann ich alles mögliche rein schreiben!“

Datentypen

Name	`20 Zeichen`
Vorname	`30 Zeichen`
Straße	`30 Zeichen`
Postleitzahl	`5 Ziffern`
Ort	`20 Zeichen`
	`20 Zeichen`

In die DB kann ich <u>nicht</u> alles reinschreiben!

Datentypen (2)

- Unterschiedliche Daten müssen unterschiedlich behandelt werden, z.B. bei
  - Sortierung
  - Rechnen

Datentypen (3)

- Beispiel Sortierung

Als Zeichen	Als Zahlen
1	1
12	2
14	12
2	14
23	23

Datentypen (4)

- Beispiel Sortierung

Als Zeichen	Als Datum
10. März 2011	04.01.2000
15.09.2013	23.02.2010
23. Februar 2010	10.03.2011
30.1.12	30.01.2012
4.1.2000	15.09.2013

Datentypen (5)

- Beispiel Rechnen
- Zahlen vs. Datum
- Konkretes Beispiel sei dem geneigten Zuhörer überlassen :-)

Datentypen (5)

- Beispiel Rechnen
- Zahlen vs. Da
- Konkretes Be
  Zuhörer über

Ohne Datenbanken:
Programmierer selbst
verantwortlich für
Einhaltung der Konsistenz
sowie Erfüllung aller
weiteren Anforderungen

Datentypen (6)

- Gibt es nicht nur bei DBMS, sondern auch in (fast) allen Programmiersprachen
- Auflistung von zur Verfügung stehenden Datentypen in DBMS folgt...

Datentypen in DBMS

`INTEGER`	ganze Zahlen
`NUMERIC(n,m), DECIMAL(n,m)`	Dezimalzahlen
`FLOAT(n), DOUBLE, REAL`	Gleitkommazahlen
`CHARACTER(n), VARCHAR(n)`	Zeichenketten
`DATE, TIME, TIMESTAMP`	Datum und Zeit
`BLOB, CLOB`	große Daten
`BOOLEAN`	wahr / falsch

wird in der nächsten Vorlesung nochmals diskutiert

RTL

Seit jeher haben die meisten ...?

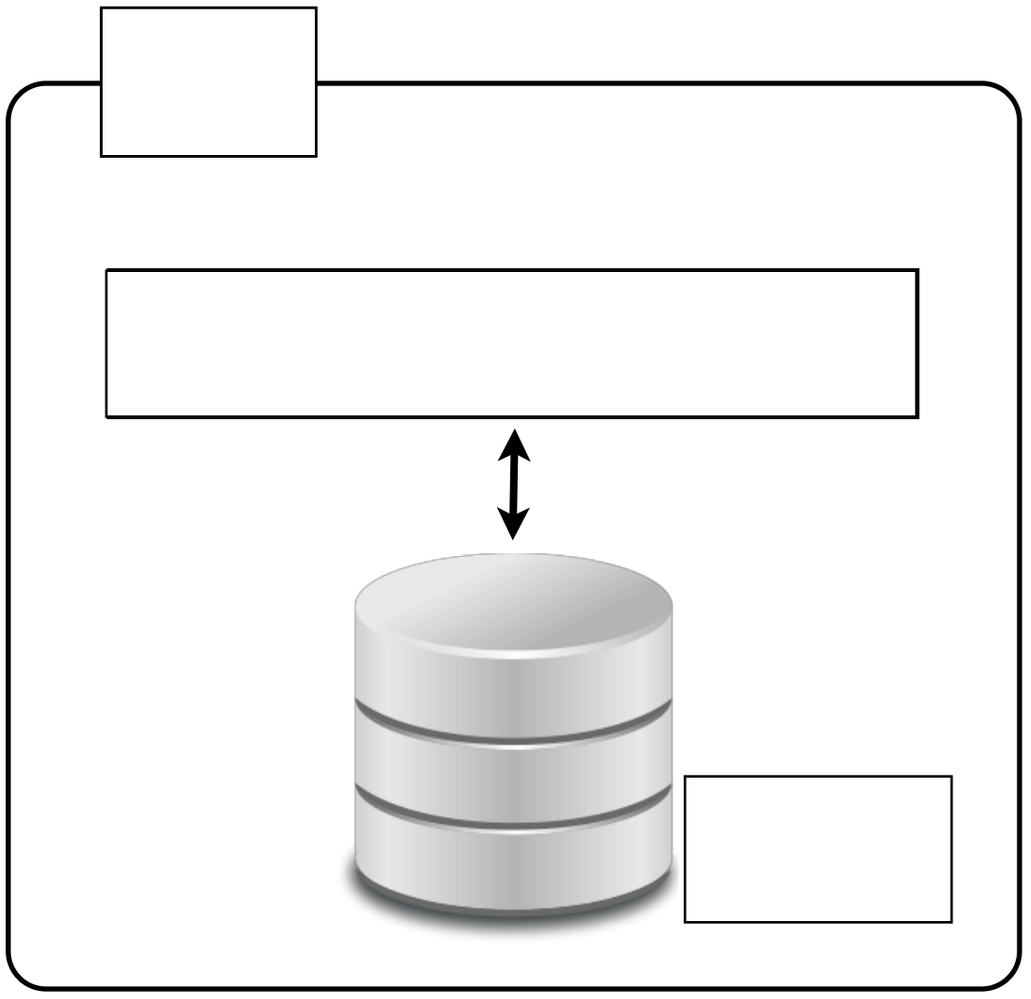
A: Dober Männer

B: Cocker Spaniels

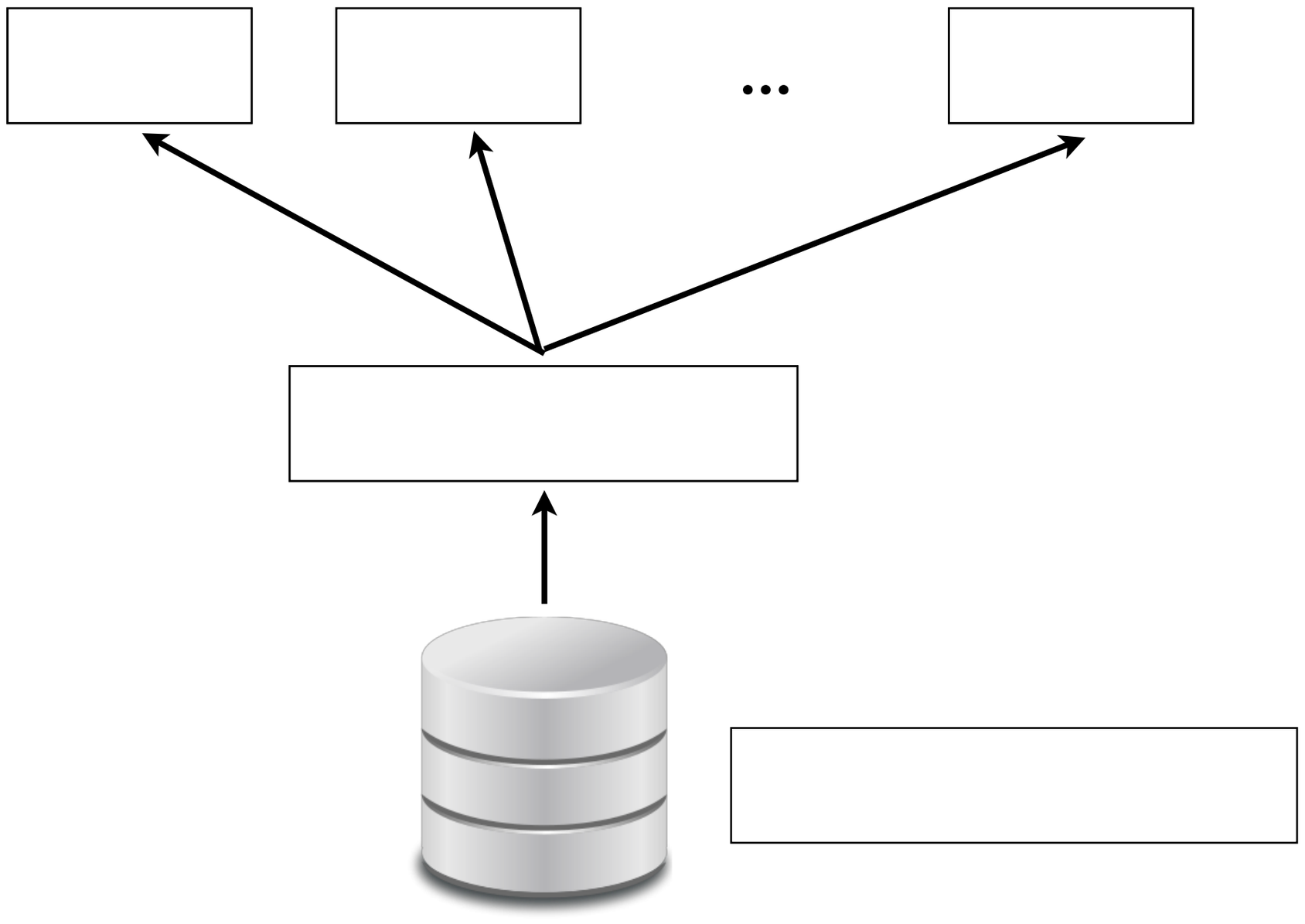
C: Schäfer Hunde

D: Riesen Schnauzer

Begriffsklärung



Datenbank-Architektur



Begriffsklärung (2)

- *Datenunabhängigkeit:*
  - Welche Arten von Datenunabhängigkeit gibt es und was bedeutet Datenunabhängigkeit?

Begriffsklärung (3)

- Was ist der Unterschied zwischen DDL und DML?
- Sie schreiben ein Anwendungsprogramm. Welches von beiden brauchen Sie...
  - in der Entwurfsphase der DB
  - im Anwendungsprogramm selbst?

Bildquellen

Film-Werbung
„Miss Daisy und
ihr Chaffeur“

„Librarian Action Figure“,
http://www.mcphee.com/laf/