

SEMINAR

# Informatik



AUFBAUKURS

# MATHEMATISCHE UND LOGISCHE GRUNDLAGEN DER DATENVERARBEITUNG

## Stellenwertsysteme

Im Laufe der geschichtlichen Entwicklung kristallisierten sich zwei verschiedene Arten der Zahlendarstellung heraus:

- zunächst **additive Zahlensysteme**: Hier wird eine Zahl durch Aneinanderfügen von Zahlzeichen dargestellt. Der Wert der Zahl ergibt sich durch Addition der Zahlzeichen (z.B. röm. Zahlen)
- später **Stellenwertsysteme**: Hier entscheidet die Stellung eines Zahlzeichens (Ziffer) innerhalb der Zahl über ihren Wert (z.B. Dezimalsystem)



## Dezimalsystem

Dieses bei uns übliche Zahlensystem hat die Zahl 10 als Basis und besitzt deshalb auch zehn verschiedene Zahlzeichen, die Ziffern: **0, 1, 2, 3, 4, 5, 6, 7, 8, 9**.

<b>T</b>	<b>H</b>	<b>Z</b>	<b>E</b>									
7	8	0	4									
				4	·	1	=	4	·	10 <sup>0</sup>	=	4
				0	·	10	=	0	·	10 <sup>1</sup>	=	00
				8	·	100	=	8	·	10 <sup>2</sup>	=	800
				7	·	1000	=	7	·	10 <sup>3</sup>	=	7000
												7804

## Andere Stellenwertsysteme

Grundsätzlich eignet sich jede Zahl, die größer als 1 ist, als Basis eines Stellenwertsystems. Das Stellenwertsystem besitzt dann stets so viele verschiedene Ziffern wie die Basis angibt. Ist die Basis  $n$ , so existieren die Ziffern  $0, 1, 2, \dots, (n-1)$ .

**Beispiel:** das Oktalsystem (Basis 8)  
**Ziffern:** 0, 1, 2, 3, 4, 5, 6, 7,  
**Stufenzahlen:**  $8^0 = 1$ ;  $8^1 = 8$ ;  $8^2 = 64$ ;  $8^3 = 512 \dots$

<b>6</b>	<b>5</b>	<b>2</b>	<b>8</b>									
				2	·	8 <sup>0</sup>	=	2	·	1	=	2
				5	·	8 <sup>1</sup>	=	5	·	8	=	40
				6	·	8 <sup>2</sup>	=	6	·	64	=	384
												426 <sub>10</sub>

## Dualsystem

Aus technischen Gründen kann ein Rechner nur zwei unterscheidbare Zustände für die Datenverarbeitung nutzen: Impuls oder Nicht-Impuls. Ein solches System heißt **Binär-** oder **Dualsystem** und besitzt nur zwei Zahlzeichen: 0 bzw. 1 (H bzw. L). Ein einzelnes solches Binärzeichen ist ein **Bit**.

Der Wertebereich einer Binärzahl wird durch die Anzahl der Ziffern als Potenz zur Basis 2 gebildet. Mit einer dreistelligen Dualzahl lassen sich Werte von 0 bis 7 darstellen. Mit einer achtstelligen Binärzahl entsprechend Werte von 0 bis 255 ( $2^8 - 1$ ) und mit einer sechzehnstelligen Binärzahl Werte von 0 bis 65535 ( $2^{16} - 1$ ). Das ist auch der Grund für den Wertebereich Integerzahlen:  $-32.768$  bis  $+32.767$ .

Das Darstellungsformat einer Fließkommazahl ist etwas komplizierter: Die meisten Software-Produkte verwenden hierfür das genormte *IEEE-Format*, das für einen einfachen genauen Wert (*Single-Datentyp*) folgenden Aufbau besitzt:

SEEEEEEE EMMMMMMM MMMMMMMM MMMMMMMM

Das Fließkommaformat besteht aus einem Bit für das Vorzeichen, acht Bit für den Exponenten und 23 Bit für die Mantisse. Durch den für den Exponenten und die Mantisse zur Verfügung stehenden Platz ergibt sich automatisch der Wertebereich und die Genauigkeit beim Rechnen mit einfach genauen Zahlen. Bei doppelt genauen Fließkommazahlen (*Double*) werden von den 64 zur Verfügung stehenden Bit 11 Bit für den Exponenten, 52 Bit für die Mantisse und 1 Bit für das Vorzeichen verwendet.

**Umwandlung Dezimalzahl – Dualzahl:**

Stufenzahlen des Dualsystems:  $2^0 = 1$                        $2^1 = 2$     $2^2 = 4$                        $2^3 = 8 \dots$

Dualzahl:	1	0	0	1	1	1	
Stufenzahlen:	$2^5$	$2^4$	$2^3$	$2^2$	$2^1$	$2^0$	
Dezimalwert:	32	-	-	4	2	1	= 39

Dabei bieten sich zwei Wege an:

- $169 = 1 \cdot 128 + 0 \cdot 64 + 1 \cdot 32 + 0 \cdot 16 + 1 \cdot 8 + 0 \cdot 4 + 0 \cdot 2 + 1 \cdot 1$   
 Reste:        41     41     9     9     1     1     1     0  
 Dualzahl:    1     0     1     0     1     0     0     1

oder übersichtlicher in Tabellenform:

169	$1 \cdot 2^7$	$0 \cdot 2^6$	$1 \cdot 2^5$	$0 \cdot 2^4$	$1 \cdot 2^3$	$0 \cdot 2^2$	$0 \cdot 2^1$	$1 \cdot 2^0$
Dualzahl	1	0	1	0	1	0	0	1

- Möglichkeit: Resteverfahren

169	:	2	=	84	Rest	1	↑
84	:	2	=	42	Rest	0	
42	:	2	=	21	Rest	0	
21	:	2	=	10	Rest	1	
10	:	2	=	5	Rest	0	
5	:	2	=	2	Rest	1	
2	:	2	=	1	Rest	0	
1	:	2	=	0	Rest	1	

Dualzahl: 1 0 1 0 1 0 0 1



REPRO MP  
 Gottfried Wilhelm Leibniz (1646-1716) entwickelte das duale Zahlensystem.

**Addition im Dualsystem:**

Rechenregeln:

$$\begin{aligned} 0 + 0 &= 0 \\ 1 + 0 &= 1 \\ 0 + 1 &= 1 \\ 1 + 1 &= 10 \end{aligned}$$

dezimal:	728	dual:	101101	=	45
	<u>+345</u>		<u>+100111</u>	=	<u>39</u>
	1073		1010100	=	84

**Subtraktion im Dualsystem:**

Die Subtraktion von Zahlen kann in jedem Zahlensystem genauso wie im Dezimalsystem durchgeführt werden. Es gibt jedoch noch eine andere Art der Subtraktion: durch **Addition der Komplementzahl**.

Auch dieses Verfahren funktioniert in allen Zahlensystemen!

*Das Komplement einer Zahl  $z$  in einem beliebigen Stellenwertsystem ist diejenige Zahl  $z^*$ , die zusammen die nächsthöhere Stufenzahl ergeben.*

*Die Komplementzahl zur Dualzahl  $z$ , die  $n$  Stellen besitzt, ist diejenige Zahl, die  $z$  zur Zweierpotenz  $z^n$  (nächste Stufenzahl) ergänzt!*

im Dezimalsystem 38       $\rightarrow$       Komplement: 62,      denn  $38 + 62 = 100$   
im Dualsystem: 1011       $\rightarrow$       Komplement: 0101,      denn  $1011 + 0101 = 10000$

Im Dualsystem gibt es einen einfachen Weg, das Komplement zu bilden:

Man ersetzt jede 0 durch eine 1 und jede 1 durch eine 0 und addiert 1:

Zahl:	1011
Zweierkomplement:	$0100 + 1 = 0101$

**Beispiel einer Subtraktion:**

10101	Minuend		
- 1011	Subtrahend	Zweierkomplement	
?		des Subtrahenden:	10100
		Addition von 1	<u>1</u>
			10101

Addition:	10101
	<u>+10101</u>
	101010

Aus mathematischen Gründen (Summe ist um  $2^n$  zu groß) tritt stets an der werthöchsten Stelle ein Übertrag auf, der zum Schluss noch zu streichen ist.

	Ergebnis	1010
Vollständige Aufgabe:	$10101 - 1011 =$	1010

### Das Sedezimalsystem bzw. Hexadezimalsystem

Da beim Übertragen von langen Ziffernfolgen im Dualsystem durchaus öfters Fehler auftreten können, hat man sich entschlossen, bei der Programmierung bzw. Übertragung in Maschinensprache ein anderes Stellenwertsystem zu verwenden: das Sedezimalsystem mit der Basis 16. Hierzu werden 16 verschiedene Zahlzeichen benötigt. Das Dezimalsystem stellt aber nur 10 verschiedene Ziffern bereit, deshalb verwendet man für die Werte zehn, elf, ..., fünfzehn die ersten 6 Buchstaben des Alphabets:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F$$

Stufenzahlen:  $16^0 = 1$        $16^1 = 16$        $16^2 = 256$        $16^3 = 4096$

Ansonsten 'funktioniert' alles genauso wie im Dualsystem bzw. in den anderen Stellenwertsystemen.

#### Beispiele für Umwandlung und Berechnung:

$$\begin{aligned} \bullet \quad 3 \text{ B } 8_{16} &= 3 \cdot 16^2 + 11 \cdot 16^1 + 8 \cdot 16^0 \\ &= 768 + 176 + 8 = 952_{10} \end{aligned}$$

$$\begin{aligned} \bullet \quad 538_{10} &= \begin{array}{r} 538 : 16 = 33 \quad \text{Rest } A \\ 33 : 16 = 2 \quad \text{Rest } 1 \\ 2 : 16 = 0 \quad \text{Rest } 2 \end{array} \\ &\text{Sedezimalzahl: } 2 \ 1 \ A_{16} \end{aligned}$$

$$\bullet \quad \text{Addition:} \quad \begin{array}{r} \text{B } 8 \\ + \text{6 A} \\ \hline 1 \ 2 \ 2 \end{array}$$

$$\bullet \quad \text{Subtraktion:} \quad 2 \ \text{F} \ \text{A} - \text{C} \ 5 = ?$$

$$\text{Komplement des Subtrahenden:} \quad \text{F} \ 3 \ \text{A} + 1 = \text{F} \ 3 \ \text{B}$$

$$\begin{array}{r} \text{Addition des Komplements:} \quad 2 \ \text{F} \ \text{A} \\ \quad \quad \quad \quad \quad \quad \quad + \ \text{F} \ 3 \ \text{B} \\ \hline 1 \ 2 \ 3 \ 5 \end{array}$$

$$\text{Streichung des Übertrags an werthöchster Stelle:} \\ \cancel{1}235 = 235$$

$$\text{Vollständige Aufgabe:} \quad 2 \ \text{F} \ \text{A}_{16} - \text{C} \ 5_{16} = 235_{16}$$

Die Umrechnung von Dualzahlen in Sedezimalzahlen ist besonders einfach, insbesondere da man in der Regel 1 Byte = 8 Bit überträgt.

$$\text{Außerdem gilt:} \quad 2^4 = 16^1 \quad 2^8 = 16^2$$

Aus diesem Grund zerlegt man 8-stellige Dualzahlen in zwei Viererblöcke, die sog. **Tetraden** und überträgt jede **Tetrade** einzeln ins Sedezimalsystem:

$$\begin{array}{c} \boxed{\phantom{0011}} \\ 1011 \quad 0110 = \mathbf{B} \quad 6 \\ \boxed{\phantom{0011}} \end{array}$$

Andere Dualzahlen, die sich nicht vollständig in Tetraden zerlegen lassen, ergänzt man zu Vierergruppen, in dem man links Nullen auffüllt:

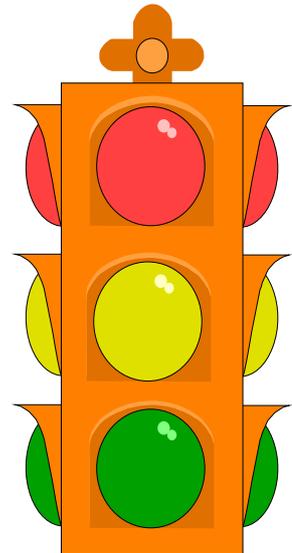
$$0011 \quad 1001 \quad 1100_2 = 3 \quad 9 \quad \mathbf{C}_{16}$$

## Logische Funktionen und binäre Schaltungen

Daten werden im Rechner durch Binärwörter dargestellt. Jede einzelne Stelle eines Binärworts ist mit einem Binärzeichen (Bit) besetzt. Aus diesem Grund kann man die Verarbeitung von Daten als Verarbeitung von Binärzeichen ansehen.

Binärzeichen stehen stellvertretend für die zwei möglichen elektrischen Zustände, die in den Bauteilen eines Computers auftreten:

Strom fließt - Strom fließt nicht  
niedrige Spannung -- hohe Spannung



Die Verarbeitung der Binärzeichen 0 und 1 geschieht mit elektrischen Schaltungen, die so aufgebaut sind, dass ihre Eingänge nur auf diese zwei verschiedenen Signale reagieren.

In einer binären Schaltung werden also die Signale 0 und 1 an den Eingängen miteinander **verknüpft** und auf ein bestimmtes Ausgangssignal **abgebildet**. Somit ist das **Ausgangssignal eine Funktion der Eingangssignale**.

Im wesentlichen lassen sich alle Schaltungen auf **3 Grundsaltungen** zurückführen:

**NOT - Schaltung**  
**AND - Schaltung**  
**OR - Schaltung**

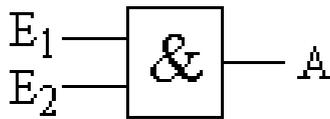
### ➤ **AND – Schaltung**

$$f(E_1, E_2) = E_1 \wedge E_2$$

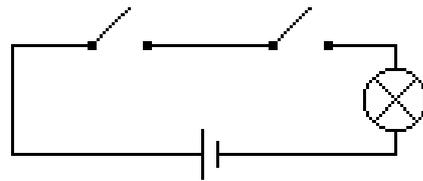
Wertetafel:

$E_1$	$E_2$	$E_1 \wedge E_2$
0	0	0
0	1	0
1	0	0
1	1	1

Schaltsymbol



Elektrische Schaltung



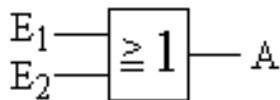
➤ OR - Schaltung

$$f(E_1, E_2) = E_1 \vee E_2$$

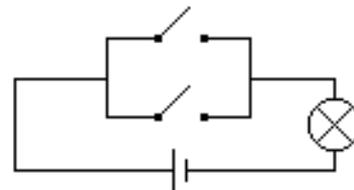
Wertetafel:

$E_1$	$E_2$	$E_1 \vee E_2$
0	0	0
0	1	1
1	0	1
1	1	1

Schaltsymbol



Elektrische Schaltung



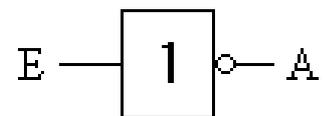
➤ NOT - Schaltung (nur 1 Eingang)

$$f(E) = \bar{E}$$

Wertetafel:

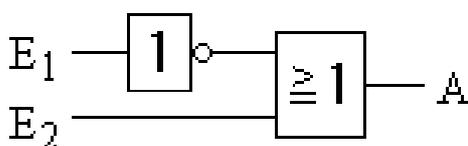
$E$	$\bar{E}$
0	1
1	0

Schaltsymbol

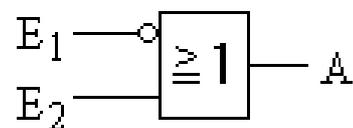


Die NOT - Schaltung kann man auch so darstellen, dass der kleine Kreis, der die Negierung darstellt, an den Eingang des unmittelbar nachfolgenden Schaltsymbols gesetzt wird.

Beispiel:  $A = E_2 \vee \bar{E}_1$



oder



Wertetafel für mehr als zwei Eingangsvariable (hier drei EV):

$$A = (E_1 \vee \overline{E_2}) \wedge E_3$$

Wertetafel	$E_1$	$E_2$	$E_3$	$\overline{E_2}$	$E_1 \vee \overline{E_2}$	$A$
	0	0	0	1	1	0
	0	0	1	1	1	1
	0	1	0	0	0	0
	0	1	1	0	0	0
	1	0	0	1	1	0
	1	0	1	1	1	1
	1	1	0	0	1	0
	1	1	1	0	1	1

### Ermittlung von Funktionsgleichungen zu vorgegebenen Wertetafeln:

Die Bezeichnungen A, E<sub>1</sub>, E<sub>2</sub> usw. sind nicht nur als Namen von elektrischen Leitungen zu verstehen, sie können auch als logische Variable betrachtet werden, die mit wahr oder falsch belegt sein können. Deswegen betrachten wir die Zeilen, deren Ausgang 'wahr' (= 1) führt und verknüpfen die Eingänge durch 'und zugleich' (  $\wedge$  ). Nun werden noch alle Fälle, die vorher mit 'wahr' bezeichnet wurden, durch 'oder auch' (  $\vee$  ) miteinander verknüpft.

$E_1$	$E_2$	$A$		
0	0	1	1	$\overline{E_1} \wedge \overline{E_2}$
0	1	1	1	$\overline{E_1} \wedge E_2$
1	0	0		
1	1	1	1	$E_1 \wedge E_2$

$$A = (\overline{E_1} \wedge \overline{E_2}) \vee (\overline{E_1} \wedge E_2) \vee (E_1 \wedge E_2)$$

Durch die Äquivalenzgesetze kann man nun die Funktionsgleichung vereinfachen:

**Kommutativgesetz:**

$$A \wedge B = B \wedge A$$

$$A \vee B = B \vee A$$

**Assoziativgesetz**

$$A \wedge (B \wedge C) = (A \wedge B) \wedge C$$

$$A \vee (B \vee C) = (A \vee B) \vee C$$

**Distributivgesetz**

$$A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$$

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$$

De Morgansche Gesetze

$$\overline{\overline{A \wedge B}} = \overline{\overline{A} \vee \overline{B}}$$

$$\overline{\overline{A \vee B}} = \overline{\overline{A} \wedge \overline{B}}$$

Anwendungsbeispiele:

$$(\overline{A \wedge B}) \vee (\overline{A \wedge B}) \vee (A \wedge \overline{B}) = \dots$$

$$\overline{A \wedge (\overline{B \vee B})} \vee (A \wedge \overline{B})$$

Distributivgesetz

$$\overline{A} \vee (A \wedge \overline{B})$$

weil  $(\overline{B \vee B})$  immer wahr

$$(\overline{A \vee A}) \wedge (\overline{A \vee B})$$

Distributivgesetz

$$\overline{(A \vee B)}$$

weil  $(\overline{A \vee A})$  immer wahr

$$\overline{A \wedge B}$$

de Morgan

$$\overline{\overline{A \wedge B}} = A \vee \overline{B}$$

de Morgan  $(\overline{\overline{A}} = A)$

**Kombination von Grundfunktionen**

In der Praxis werden neben den Verknüpfungen NOT, AND, OR noch weitere Schaltungen verwendet:

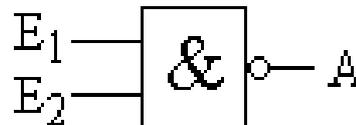
> **NAND-Schaltung**

$$f(E_1, E_2) = \overline{E_1 \wedge E_2}$$

Wertetafel

E <sub>1</sub>	E <sub>2</sub>	E <sub>1</sub> ∧ E <sub>2</sub>	A
0	0	0	1
0	1	0	1
1	0	0	1
1	1	1	0

Schaltsymbol



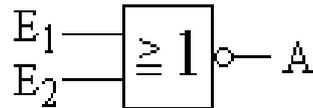
> **NOR-Schaltung**

$$f(E_1, E_2) = \overline{E_1 \vee E_2}$$

Wertetafel

E <sub>1</sub>	E <sub>2</sub>	E <sub>1</sub> ∨ E <sub>2</sub>	A
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0

Schaltsymbol

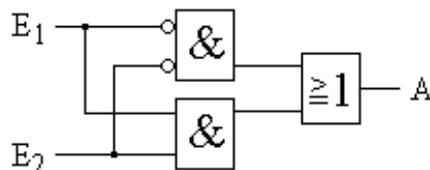
Äquivalenzschaltung

Hierbei handelt es sich um eine Schaltung, die zwei Binärstellen auf Gleichheit überprüft und genau dann ein 1-Signal liefert.

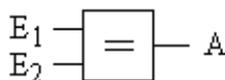
$E_1$	$E_2$	$A$		
0	0	1	1	$\overline{E_1} \wedge \overline{E_2}$
0	1	0		
1	0	0	1	$E_1 \wedge E_2$
1	1	1		

$$f(E_1, E_2) = \underline{(\overline{E_1} \wedge \overline{E_2}) \vee (E_1 \wedge E_2)}$$

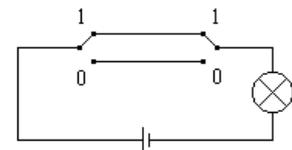
Ausführliche Schaltung:



Schaltsymbol:

**Beispiel für eine weitere Anwendung der Äquivalenzschaltung:**

Hier fließt genau dann Strom, wenn beide Schalter auf 0 oder beide Schalter auf 1 stehen! Wechselschaltung:

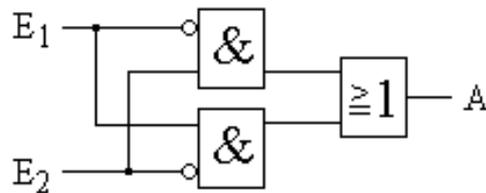
Antivalenzschaltung:

Gegenteil der Äquivalenzschaltung: sie liefert am Ausgang genau dann ein 1-Signal, wenn beide Eingänge unterschiedlich geschaltet sind:

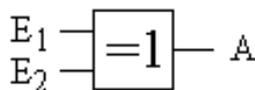
$E_1$	$E_2$	$A$		
0	0	0	1	$\overline{E_1} \wedge E_2$
0	1	1		
1	0	1	1	$E_1 \wedge \overline{E_2}$
1	1	0		

$$f(E_1, E_2) = \overline{(E_1 \wedge E_2)} \vee (E_1 \wedge \overline{E_2})$$

Ausführliche Schaltung:



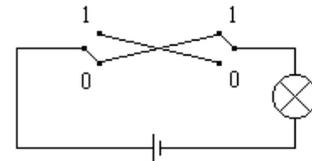
Schaltsymbol:



Bei einer 'verdrehten' Wechselschaltung kommt die Antivalenzschaltung zur Anwendung:

Hier fließt genau dann Strom, wenn jeweils ein Schalter auf 0 und gleichzeitig der andere auf 1 steht.

'verdrehte'  
Wechselschaltung:

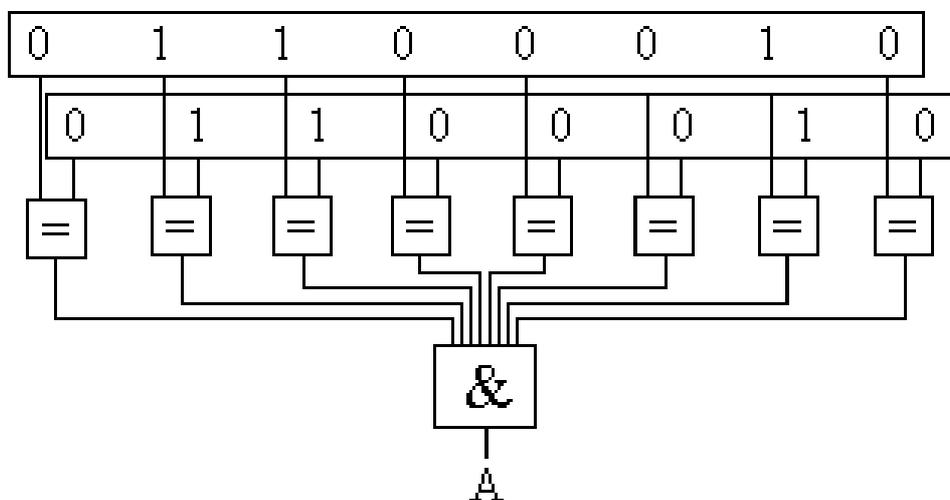


## Vergleicher

Ein Vergleicher ist eine Schaltung, die zwei Binärwörter (z.B. Bytes) bitweise auf ihre Gleichheit hin überprüft.

Sind die Binärwörter gleich, so liefert die Schaltung am Ausgang ein 1-Signal, bei Ungleichheit ein 0-Signal.

Für den bitweisen Vergleich eignet sich die Äquivalenzschaltung. Da bei einem Byte 8 Bit auf Gleichheit hin untersucht werden müssen, benötigt man 8 Äquivalenzschaltungen, die ihrerseits über eine AND-Schaltung miteinander verknüpft werden, da ja nur dann am Ausgang ein 1-Signal erscheinen soll, wenn alle acht Stellen übereinstimmen.



## Addition von Dualzahlen im Computer

Für den Computer (=Rechner) ist die Addition von einzelnen Bit von grundlegender Bedeutung, denn

- kann man zwei einstellige Dualzahlen addieren, so können auch mehrstellige Dualzahlen addiert werden.
- wenn addiert werden kann, so kann auch multipliziert werden, da die Multiplikation ja nur eine verkürzte Schreibweise für eine Addition von lauter gleichen Summanden ist.
- kann multipliziert werden, so ist auch das Potenzieren kein Problem.
- wenn addiert werden kann, so kann auch subtrahiert (Subtraktionsersatz) werden und somit kann auch dividiert werden.

Für die Division gilt, dass sie nicht nur als Ganzzahldivision mit Rest funktioniert, sondern auch für die 'normale' Division:

$$\begin{array}{r}
 7 : 4 = \underline{1} \\
 \text{Rest } 3 \\
 \text{Rest } 3 \text{ mal } 10 = 30; \\
 \text{Rest } 2 \text{ mal } 10 = 20;
 \end{array}
 \qquad
 \begin{array}{r}
 30 : 4 = \underline{7} \text{ Rest } 2 \\
 20 : 4 = \underline{5} \text{ Rest } 0
 \end{array}$$

also  $7 : 4 = 1,75$

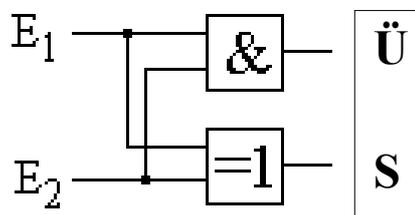
### Addition von 2 Bit (Halbaddierer)

$E_1$	$E_2$	$\ddot{U}$	S
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

Für die Spaltsumme gilt:  $S = (\overline{E_1} \wedge E_2) \vee (E_1 \wedge \overline{E_2})$

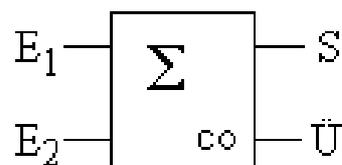
Für die Übertragungsspalte gilt:  $\ddot{U} = E_1 \wedge E_2$

Somit ergibt sich folgende Schaltung:



Schaltsymbol des Halbaddierers

$\Sigma$  = Summenzeichen  
co = carry output

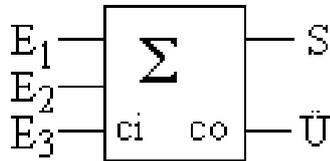


### Addition von 3 Bit (Volladdierer)

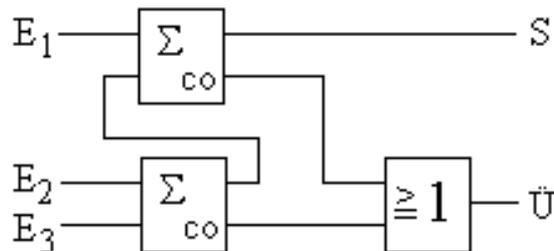
Sollen drei einstellige Dualzahlen addiert werden (Übertrag von vorher) reicht ein Halbaddierer nicht mehr aus. Es werden drei Eingänge benötigt.

Schaltsymbol:

ci = carry input

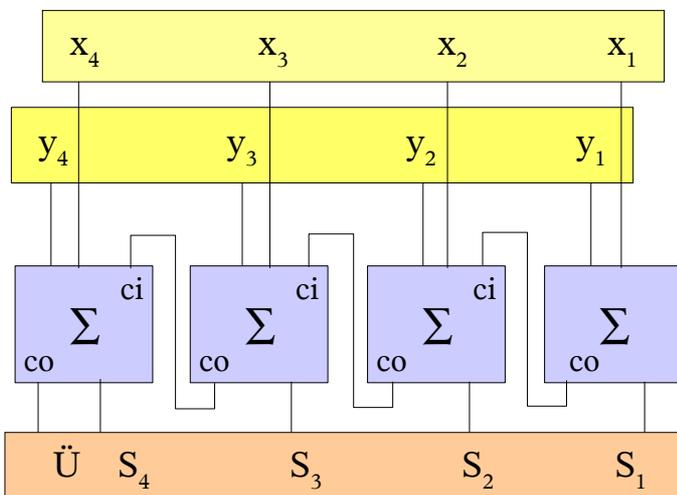


Ausführliches Schaltbild (2 Halbaddierer und OR-Schaltung):



E <sub>1</sub>	E <sub>2</sub>	E <sub>3</sub>	Ü	S
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	0	1
1	0	1	1	0
1	1	0	1	0
1	1	1	1	1

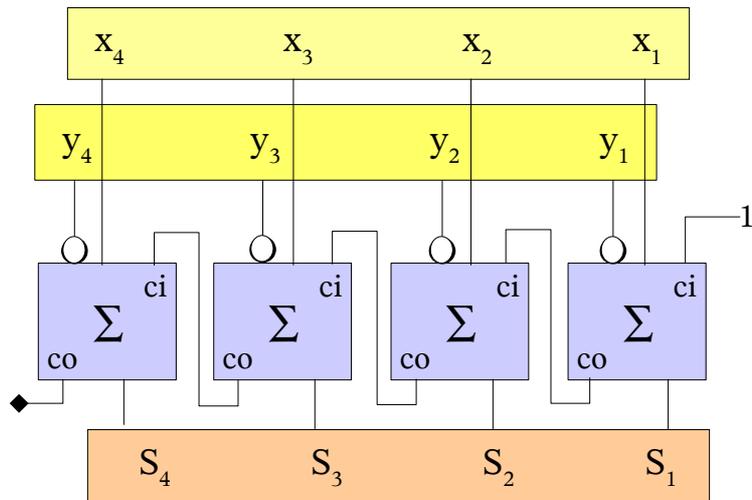
### Addition von mehrstelligen Dualzahlen



Hier verwendet man eine Schaltung mit mehreren Volladdierern und einen Halbaddierer, der für die erste Stelle ganz rechts verwendet wird.

## Subtraktion von mehrstelligen Dualzahlen

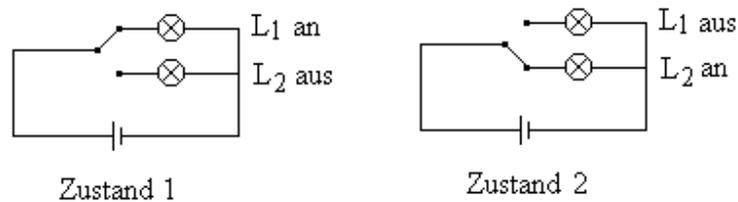
Wie bereits oben gezeigt erfolgt die Subtraktion über die Addition des Komplements. Entsprechend kann die Schaltung abgeändert werden.



## Das RS-Flipflop

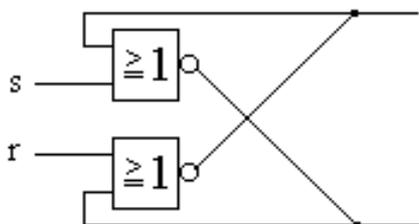
Alle bisher aufgeführten Digitalschaltelemente können ihren jeweiligen Zustand an den Ausgängen nur solange aufrecht erhalten, wie ein Takt des Computers dauert. Da aber das Speichern von Information im Computer eine sehr wichtige Aufgabe ist, konstruiert man sog. Kippschaltungen, die ihren Zustand solange beibehalten, bis über eine besondere Eingangsleitung ein neuer Zustand gesetzt wird. Solche Speicherelemente nennt man Flipflops. Sie können jeweils 1 Bit an Information speichern. Durch eine Aneinanderreihung vieler solcher Flipflops lassen sich große Computerspeicher realisieren.

Das Verhalten solcher Flipflopschaltungen ist mit einem Umschalter vergleichbar, der zwei Lampen an- und ausschaltet.

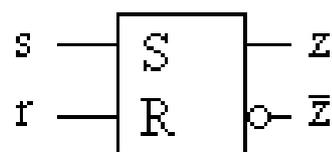


Man erkennt, dass der Umschalter nicht beide Positionen gleichzeitig einnehmen kann. Genauso ist es unmöglich, den Zustand nur einer einzigen Lampe zu verändern.

Ein Digitalbaustein mit diesen Eigenschaften muss also zwei Eingangsleitungen zum Setzen  $s$  und zum Rücksetzen  $r$  aufweisen, sowie zwei Ausgangsleitungen  $z$  und  $\bar{z}$ . Für den Aufbau werden NOR-Gatter verwendet.



Schaltbild des RS-Flipflops:



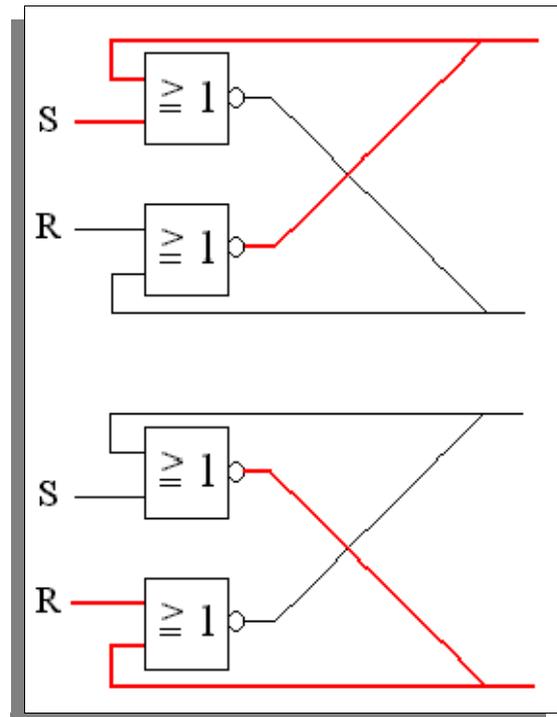
Um den Zustand zu setzen oder zurückzusetzen reicht ein kurzer Stromimpuls am Eingang  $r$  bzw.  $s$  eines Flipflops aus. Trotzdem wird an den Ausgängen stets ein Dauerzustand ausgegeben.

### Flipflop setzen:

- Ein 1-Signal am S-Eingang von NOR<sub>1</sub> bewirkt eine 0 an seinem Ausgang
- Diese 0 überträgt sich auf den Eingang von NOR<sub>2</sub>
- Beide NOR<sub>2</sub>-Eingänge führen nun 0, was zu einem 1-Signal an seinem Ausgang führt
- Diese 1 wiederum überträgt sich auf NOR<sub>1</sub>, ändert aber dessen Zustand nicht

### Flipflop rücksetzen:

- Ein 1-Signal am R-Eingang von NOR<sub>2</sub> bewirkt eine 0 an seinem Ausgang
- Diese 0 überträgt sich auf den Eingang von NOR<sub>1</sub>
- Beide NOR<sub>1</sub>-Eingänge führen nun 0, was zu einem 1-Signal an seinem Ausgang führt
- Diese 1 wiederum überträgt sich auf NOR<sub>2</sub>, ändert aber dessen Zustand nicht



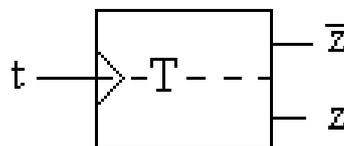
Ein Impuls an **S** *setzt* das Flipflop, d.h. Ausgang 1 geht auf 1, Ausgang 2 geht auf 0

Ein Impuls an **R** *rücksetzt* das Flipflop, d.h. Ausgang 1 geht auf 0, Ausgang 2 geht auf 1

### Dualzähler

Das RS-Flipflop hat einen großen Nachteil: Die Schaltung im Computer muss dafür sorgen, dass **kein undefinierter Zustand** eintritt. Es darf hier nur die r-Eingangslleitung oder nur die s-Eingangslleitung mit einem Stromimpuls versorgt werden. Um diesen Nachteil zu umgehen, entwickelte man das T-Flipflop. Es besitzt nur einen Takteingang, bei dem jeder Impuls auf diesen Eingang bewirkt, dass beide Ausgangslleitungen ihren Zustand ändern.

Schaltbild eines T-Flipflops:

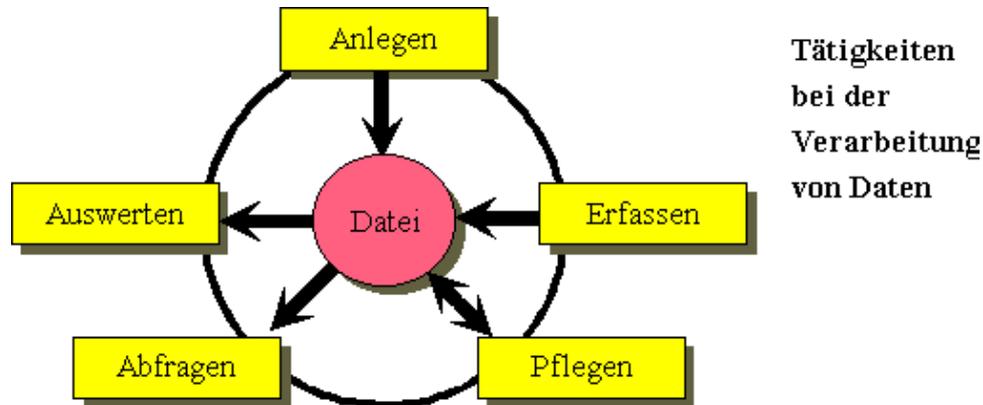


Seine Funktionsweise lässt sich mit Tastschaltern vergleichen, die z. B. in Treppenhäusern verwendet werden: Drückt man die Taste einmal, geht das Licht an, drückt man erneut, geht das Licht wieder aus.

Durch das Hintereinanderschalten solcher T-Flipflops lassen sich sog. **Zählerschaltungen** aufbauen. Pro eintreffendem Impuls wird um eine Stufe weiter gezählt. Verwendet man beispielsweise vier T-Flipflops, so kann der Zähler bis 15 zählen.

## DATENBANKEN

Bei der (elektronischen) Datenverarbeitung handelt es sich in der Praxis um immer wiederkehrende Tätigkeiten:



Damit bei dieser Arbeit möglichst wenige Probleme auftauchen, sind einige Vorüberlegungen erforderlich.

### AUFBAU VON DATENBANKEN (von M. Michelson)

#### Was ist eine Datenbank?

Datensammlungen mit herkömmlicher Verwaltung erschließen sich dem Benutzer nur sehr beschränkt: Sie sind, ob als Karteien oder in Ordnern angelegt, jeweils nach einem Kriterium geordnet. Soll Zugriff unter mehreren Aspekten auf die Datensammlung gewährleistet sein, so muss sie nach dem herkömmlichen Prinzip mehrmals identisch vorhanden sein. So könnte eine Adressliste einmal nach Namen, dann nach Wohnorten und nach Berufen der enthaltenen Personen geordnet sein. In wissenschaftlichen Bibliotheken sind immer noch drei Kataloge die Regel: Der Autorenkatalog, der systematische Katalog und der Schlagwortkatalog. Informationen, nach denen nicht geordnet wurde, lassen sich überhaupt nicht suchen. So ist keinem herkömmlichen Bibliothekskatalog zu entnehmen, welche Bücher eines bestimmten Verlages im Bestand sind.

Datenbanken sollen nun diese Einschränkungen in der konventionellen Datenhaltung überwinden. Im allgemeinen Sinne versteht man unter einer Datenbank eine elektronisch organisierte Datensammlung zu einem thematisch oder organisatorisch umgrenzten Gebiet. Im engeren Sinne bezeichnet eine Datenbank nur den Datenbestand in einem **Datenbanksystem**. In diesem engeren Sinne besteht das Datenbanksystem aus der Datenbank, dem Datenbankverwaltungssystem (auch: Datenbankmanagementsystem DBMS; Datenbankprogramm; häufig auch selbst Datenbanksystem genannt) und der Benutzeroberfläche oder den Anwendungsprogrammen. In der Regel kann ein Datenbankverwaltungssystem auch mit unabhängigen Anwendungsprogrammen kommunizieren, wenn eine entsprechende Programmschnittstelle eingerichtet ist.

Während ein **Anwendungsprogramm** ein eigenständiges Programm mit Verarbeitungsfunktionen ist, dient die **Benutzeroberfläche** allein der Kommunikation des Nutzers mit dem Datenbankverwaltungssystem. Üblich ist eine spezielle Befehlssprache, ein menügeführter Dialog oder eine Eingabemaske.

Das **Datenbankverwaltungssystem** ist die Software, die die Erstellung, die Verwaltung und den Zugriff auf eine Datenbank ermöglicht. Sie sorgt für ein heute entscheidendes Kriterium in der Organisation eines Datenbanksystems: Der logischen und physischen Trennung von Datenbank und

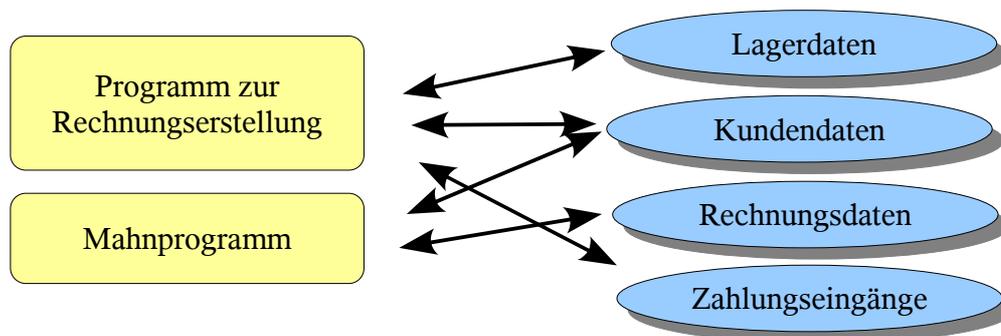
Anwendungsprogrammen. Diese Trennung lässt bei Änderungen in den Daten die Anwendungsebenen unberührt und umgekehrt haben Änderungen bei den Programmen keinen Einfluss auf die Daten. Das Datenbankverwaltungssystem fungiert praktisch als Zwischenglied zwischen den internen Daten und der externen Benutzerebene.

Diese Trennung ist normalerweise nur bei großen Datenbanksystemen, die auf zentralen Rechnern installiert sind, von Relevanz. Hier können die verschiedensten Nutzer mit der Datenbank arbeiten und zwar je nach ihrem Bedarf mit unterschiedlichen Anwendungsprogrammen. Bei den marktgängigen PC-Datenbankprogrammen wie dBase bietet sich das Datenbanksystem als eine Einheit dar.

In einer Datenbank sollen alle Informationen nur einmal gespeichert sein (Redundanzfreiheit), um bei einer Änderung keine Inkonsistenzen zu bekommen. Eine Ausnahme bilden **Schlüsselmerkmale**, die als identifizierendes und verknüpfendes Element in mehreren Dateien enthalten sein können.

### Datenbanksystem und Datenmodelle

Die einfachste Möglichkeit ist, mit Hilfe einzelner Dateiverwaltungsprogramme auf die Daten der **Datenbasis** (das ist die Gesamtheit aller Daten) zuzugreifen:



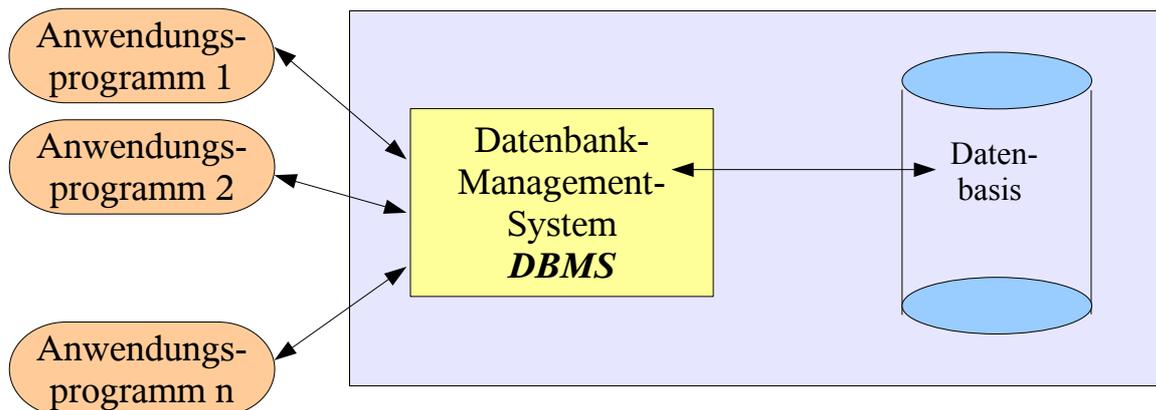
Dieses einfache und deshalb weit verbreitete Prinzip hat mehrere Nachteile:

- viele Daten (z. B. Personennamen) können in mehreren Dateien gespeichert sein (Redundanz)
- großer Speicherplatzbedarf
- werden Änderungen nicht in allen Dateien durchgeführt, so führt das zu Widersprüchen (Im Bereich der DV/Logik versteht man unter **Konsistenz die Widerspruchsfreiheit**.)

Für einen Arbeitsauftrag werden oft nicht alle Datenfelder eines Datensatzes gebraucht. Trotzdem muss das verarbeitende Programme den ganzen Record berücksichtigen. Dies kann zu Datenschutzproblemen führen.

- Schutzmaßnahmen müssen in jedem einzelnen Programm vorgenommen werden
- jedes eingesetzte Programm muss den Aufbau der Datei (Organisationsform und Struktur der Datensätze) kennen.
- wird die Struktur eines Datensatzes geändert, müssen alle Programme, die darauf zugreifen, ebenfalls geändert werden: **Programme sind von den Daten abhängig**

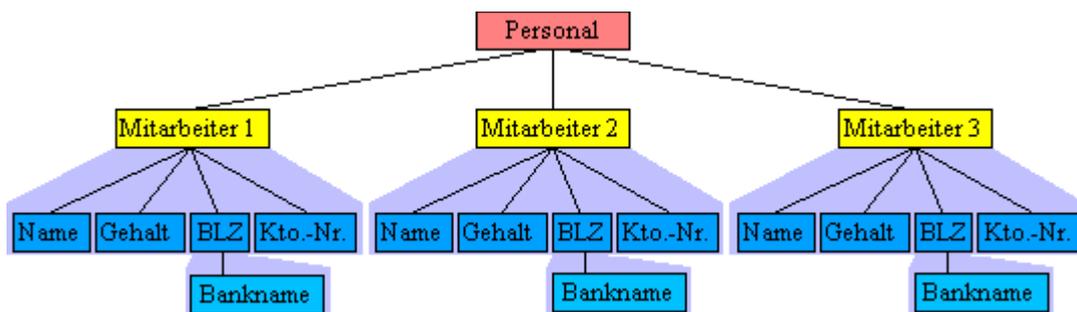
Diese Nachteile lassen sich durch den Einsatz eines Datenbankmanagementsystems weitgehend vermeiden. Die Anwenderprogramme greifen hier über das DBMS (= System von Verwaltungs- und Kontrollprogrammen) auf die Datenbasis zu. Die Datenbasis kann auf mehrere Rechner verteilt sein.



### Datenmodelle

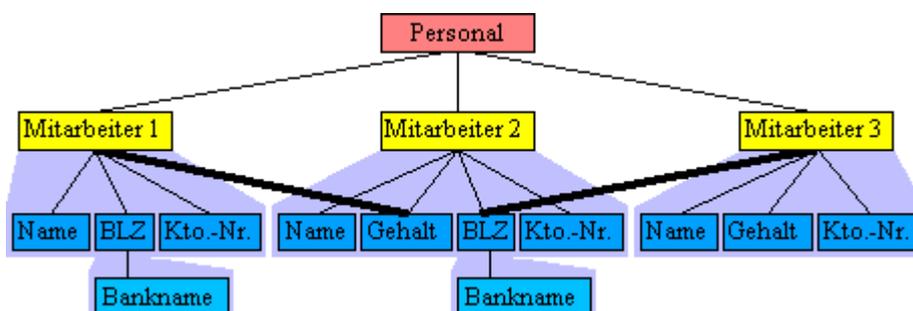
Zum Entwurf einer Datenbank existieren verschiedene **Datenmodelle**:

Das **hierarchische Modell** ist baumorientiert. Der Zugang zu einem Datensatz erfolgt über eine Wurzel, z.B. über die *Mitarbeiternummer*. Alle mit der Wurzel verknüpften Daten können nur über die Knoten erreicht werden.



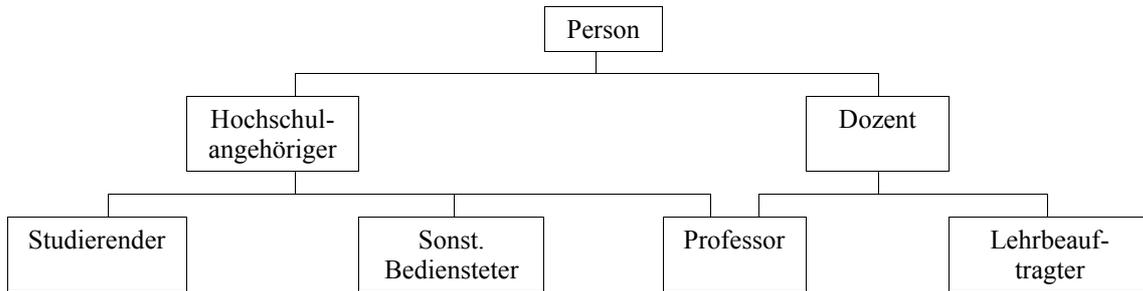
Der Nachteil dieses Modells besteht in seiner Unflexibilität. Änderungen an der Struktur führen oft zu einer verwirrenden Gesamtstruktur. Das Hierarchiemodell wird heute kaum noch verwendet.

Das **netzwerkartige Datenmodell** ist ebenfalls baumorientiert, läßt jedoch die Verknüpfung von Knoten über verschiedene Ebenen zu.



Das **objektorientierte Datenmodell** enthält als Basiselemente **Objekte**, **Eigenschaften (Literele)** und **Methoden**. Eine objektorientierte Datenbank speichert Objekte. Sie basiert auf einem Schema und enthält Instanzen, die sich auf den im Schema definierten Typen gründen. Die Typen dienen zur abstrakten Beschreibung von Objekten hinsichtlich ihrer Zustände und ihres Verhaltens. Typen können auch vom Benutzer definiert werden (vergleichbar mit dem Verbund). Außerdem kennt dieses Datenmodell die Vererbung, den Polymorphismus und die Kapselung. Durch das Vererb-

ungskonzept wird auch die Hierarchiebildung unterstützt. Die Vererbung kann sowohl auf Strukturebene als auch auf Verhaltensebene erfolgen. Für eine Hochschule beispielsweise wäre folgende Struktur denkbar:

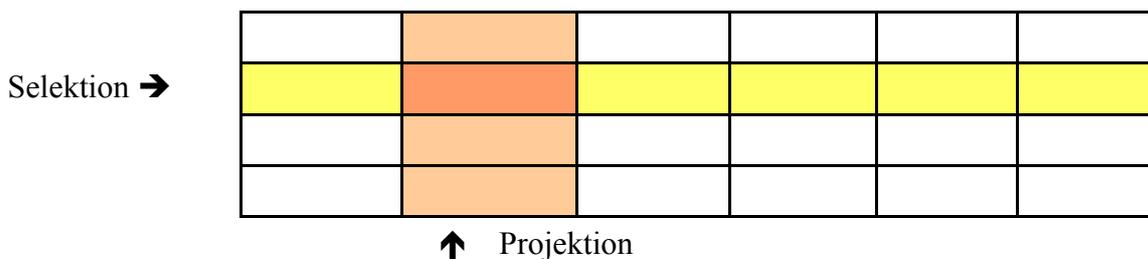


Das **relationale Datenmodell** wird gegenwärtig am häufigsten verwendet. Sein Ursprung liegt wohl in dem Artikel „*A Relational Model of Data for Large Shared Data Banks*“ von E. F. Codd, der bereits 1970 in Amerika erschien. Es verknüpft Tabellen, die in Beziehung zueinander stehen (Relationen) und die Daten in Zeilen und Spalten enthalten. Ihr Aufbau folgt definierten logischen Regeln. Mehrere Tabellen sind untereinander ebenfalls durch logische Verknüpfungen verbunden.

Personal-Nr.	Name	Gehalt	Kto.Nr.	BLZ	BLZ	Bankname
8741	Maier	2896,45	175634	76020050	76020050	A-Bank
2645	Huber	2578,34	176289	76020050	84010010	B-Bank
7856	Hobel	3189,51	179988	84010010		

Kennzeichen des relationalen Datenmodells:

- alle Daten werden in Tabellen als Werte von Relationen dargestellt
- der Zugriff auf die Daten erfolgt über die Datenfeldinhalte
- es gibt die Operationen Selektion, Projektion und Join (Verbinden)



**Anlegen einer Tabelle**

Zunächst muss in jedem Fall eine neue Datei (Tabelle) erzeugt werden. Dabei wird bei manchen Systemen als erstes der Name der Tabelle abgefragt und anschließend die Struktur der Tabelle festgelegt. Für jedes Merkmal (Datenfeld, Attribut) sind folgende Angaben zu machen:

**FELDDNAME                      FELDTYP                      LÄNGE                      INDEX**

Der **FELDDNAME** muss mit einem Buchstaben beginnen und kann auch Ziffern und Unterstreichungszeichen enthalten.

Folgende **FELDTYPEN** findet man in praktisch allen Datenbanksystemen:

- alphanumerischer Datentyp (Zeichenkette)

- numerischer Datentyp (Zahlen)
- Datum-Datentyp
- logischer Datentyp
- Memo-Datentyp

Nun soll dies an einem Beispiel mit einer möglichst einfachen Tabelle realisiert werden. Die Tabelle besteht aus Holzklötzen, für die nur die Merkmale Name, Form, Farbe und Stellplatz in einem Regalsystem angelegt werden. Für das Gewicht beispielsweise interessieren wir uns nicht. Nachdem wir die vier Datenfelder für unsere Beispieldatei angelegt haben, können wir bereits die ersten Attributwerte eingeben und somit Datensätze erzeugen.

 12	 15	 1	 5
 20	 2	 11	 4
 13	 3	 8	 17
 19	 14	 6	 7
 18	 15	 10	 9

In dieser Datei werden vier verschiedene Attribute verwendet.

<u>Name</u>	<u>Form</u>	<u>Farbe</u>	<u>Platz</u>
A	Kreis	rot	1
B	Dreieck	blau	2
C	Viereck	gelb	3
D		grün	4
:			:
T	Daraus ergeben sich 4800 mögliche Kombinationen. Nur wenige werden verwendet.		20

Derartige Dateien lassen sich grundsätzlich folgendermaßen in *Tabellenform* darstellen:

A T T R I B U T E				
	NAME	FORM	FARBE	PLATZNR
1	D a	t e n	s a	t z
2	D a	t e n	s a	t z
3	D a	t e n	s a	t z
4	D a	t e n	s a	t z
:				Datenfeld
20	D a	t e n	s a	t z

↑  
Projektion

In einer einzelnen Datei existieren nur zwei grundlegende Bearbeitungsmöglichkeiten:

### Selektion und Projektion

Durch **Projektion** werden **einzelne Attribute (Merkmale)** ausgewählt.

Bei einer **Selektion** werden **ein oder mehrere Datensätze** ausgewählt.

Jedes einzelne Element dieser Objektmenge muss bei der elektronischen Verarbeitung in einer Datenbank eindeutig identifizierbar sein. Jeder Mensch gehört zu einer bestimmten Gruppe (Männer oder Frauen, zu einer bestimmten Berufsgruppe, Autofahrer oder Bahnfahrer ...). Innerhalb einer solchen Gruppe gibt es bestimmte Gemeinsamkeiten, andererseits aber lässt sich auch jedes Element auf Grund individueller Unterschiede herausfiltern und damit identifizieren. Dabei muss das Identifizierungsmerkmal nicht unbedingt mit den üblichen Gruppeneigenschaften zu tun haben, z.B. ist jeder Mensch auf Grund seiner Fingerabdrücke aus einer Gruppe herausfinden.

Existiert in einer Datei kein eindeutiges Identifizierungsmerkmal, so kann man ein solches zusätzlich definieren. Oft sind dies Nummern, die einem bestimmten Schema unterliegen (Artikelnummern, Rechnungsnummern, ISBN) oder im einfachsten Fall ein Zähler. In Datenbanktabellen sind solche Felder zur eindeutigen Identifizierung eines Datensatzes von elementarer Wichtigkeit. Man spricht von einem **Schlüsselfeld, Primärschlüssel oder nur Schlüssel**. Ein Primärschlüssel kann auch aus mehreren Eigenschaften (Feldern) zusammengesetzt sein.

Ein Schlüsselement darf nur auf einen Datensatz zutreffen. Ein möglicher Schlüssel wäre im obigen Beispiel das Datenfeld **'Stellplatz'**.

Die einzelnen Datenfelder können auch durch einen **Index (Sortierschlüssel)** gekennzeichnet werden. Im Index wird die Position der Datensätze gespeichert. In der Regel indiziert man die Datenfelder, in denen häufig gesucht oder sortiert wird, da die Suche und das Sortieren in Indexdateien sehr viel schneller abläuft. Das Primärschlüsselfeld wird automatisch indiziert! Die verschiedenen DBMS verwenden unterschiedliche Möglichkeiten der Speicherung von Indexdateien. In Access werden sie intern gespeichert, andere, wie z.B. dBase for Windows, speichern sie extern.

Neben den grundlegenden Objekten der Datenbasis einer Datenbank wie Datenfeld, Datensatz und Datei gibt es auch Objekte wie Tabelle, Abfrage, Formular, Bericht, Makro, und Modul. Diese Objekte ermöglichen verschiedene Funktionen des DBMS (Datenbankmanagementsystems):

- **Tabelle** – Erstellen der Datensatzstruktur, Eingabe der Daten
- **Formular** – Ein- und Ausgabe der Daten in Maskenform
- **Abfrage** – Suchen, Sortieren, Berechnen und Filtern von Daten
- **Bericht** – Ausgabe der Daten (vorzugsweise auf Drucker) ggf. mit Berechnungen
- **Makro, Modul** – Steuern von Programmabläufen

Ein **Formular** ist nichts anderes als eine Bildschirmmaske, die der Ein- bzw. Ausgabe der Daten dient. Auch die bereits erzeugten Abfragen können mit Formularen am Bildschirm dargestellt werden. Die einzelnen Datenbankprogramme besitzen meist einen Assistenten zum Erstellen standardisierter Formulare. Individuelle Anpassungen sind möglich.

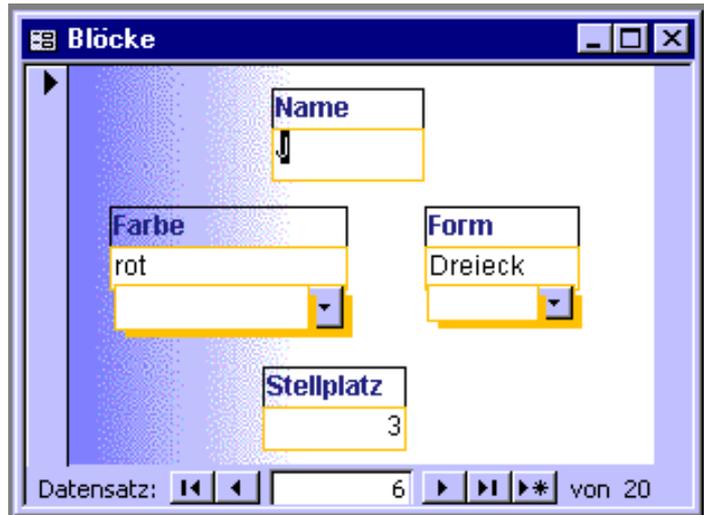
### *Blöcke*

<i>Stellplatz</i>	<i>Name</i>	<i>Form</i>	<i>Farbe</i>
1	F	Viereck	gelb
2	O	Viereck	gelb
3	J	Dreieck	rot
4	E	Dreieck	rot
5	R	Kreis	grün
6	N	Dreieck	gelb
7	C	Kreis	rot
8	Q	Dreieck	rot
9	A	Dreieck	grün
10	P	Viereck	grün
11	K	Viereck	blau

Berichte sind vorzugsweise zum Ausdruck bestimmt. Auch hierfür steht ein Assistent zur Verfügung.

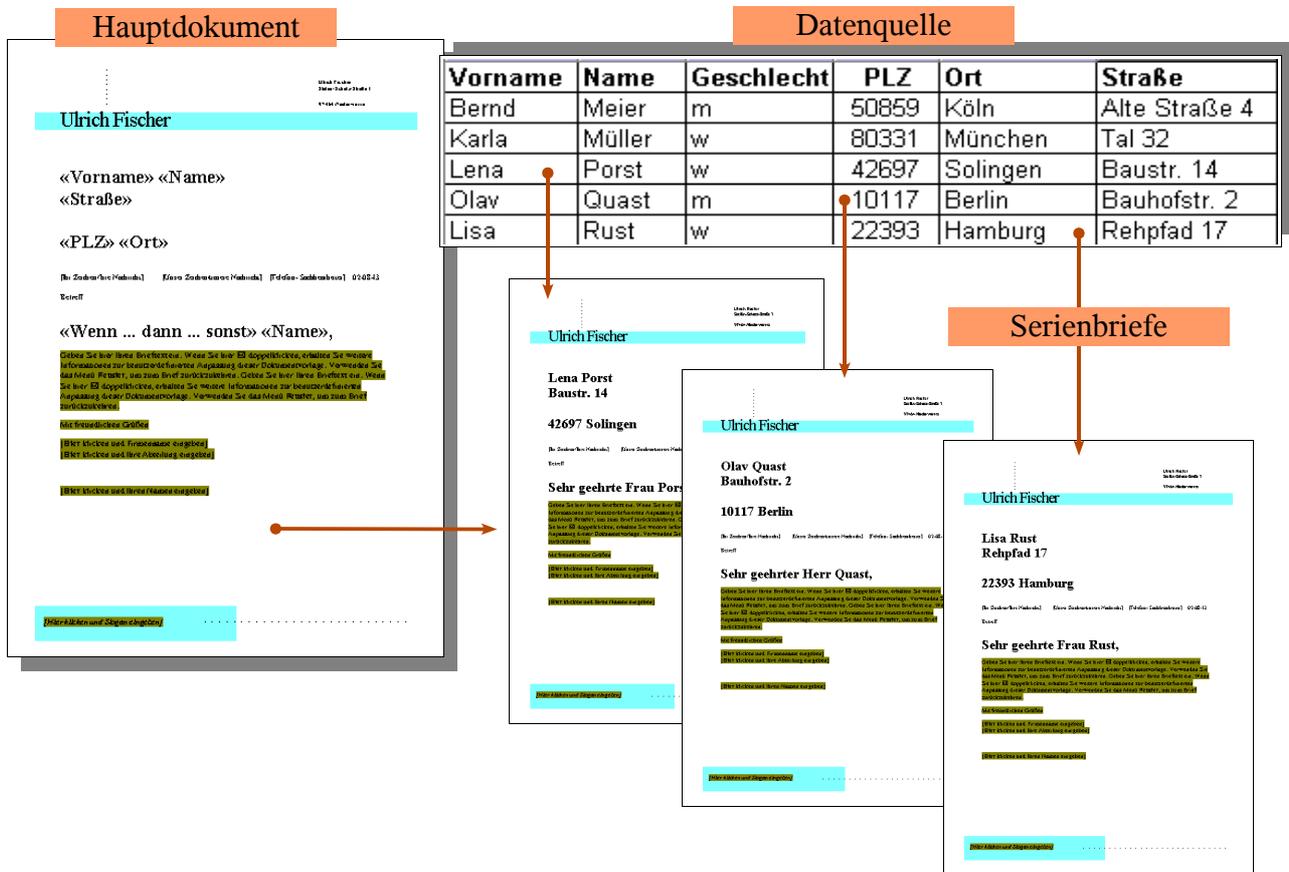
Sicher am häufigsten arbeitet man in einem Datenbankprogramm mit dem Objekt 'Abfrage'. Sinn und Zweck von Abfragen sind:

- x Daten auf unterschiedliche Arten anzuzeigen
- x Daten zu analysieren
- x Daten zu verändern
- x Daten für Formulare und Berichte aufzubereiten
- x neue Daten aus vorhandenen Daten berechnen



Auch hier erleichtern Assistenten den ersten Umgang mit dieser Funktionalität.

Eine der häufigsten Anwendungen im Zusammenhang mit Datenbanken ist der Serienbrief. Dabei handelt es sich um ein Dokument, das mit gleichem Hauptinhalt, aber individuellen Angaben, z.B. Adresse, Anrede usw., an verschiedene Personen gerichtet ist. Wichtig dabei ist die Unterscheidung zwischen der Steuerdatei (Datenquelle) und dem Hauptdokument. Beim Seriendruck werden dann beide Dateien zusammengeführt und eine festgelegte Anzahl von Serienbriefen erzeugt. Nachfolgende Grafik soll dies veranschaulichen.



Alle bisher beschriebenen Leistungsmerkmale eines solchen relationalen Datenbanksystems ließen sich auch mit einem Tabellenkalkulationsprogramm umsetzen, da Grundlage für alle Darstellungen die Tabellenform ist. Die eigentliche Stärke von Datenbanken liegt in der Möglichkeit, Daten aus mehreren Tabellen miteinander zu verknüpfen. Beispielsweise ist es notwendig, in einer Firma die Tabelle 'Aufträge' mit der Tabelle 'Kunden' zu verknüpfen. Das funktioniert jedoch nur, wenn in beiden Tabellen ein gemeinsames Schlüsselfeld vorhanden ist.

Die Leistung einer Datenbank lässt sich optimieren, wenn man folgende Richtlinien befolgt:

- Tabellen sollen keine redundanten Daten enthalten. (Redundanz = Mehrfachspeicherung von Daten)
- Für die Datenfelder muss der richtige Datentyp und die optimale Feldgröße festgelegt werden.
- Felder, die sortiert, mit anderen verknüpft oder für die häufig Kriterien festgelegt werden sollen, sind zu indizieren (Sortierschlüssel festzulegen), um bei großen Datenbeständen die Arbeitsabläufe zu beschleunigen.
- Indexdateien benötigen Speicherplatz und verlangsamen das Hinzufügen, Löschen oder Aktualisieren des Datenbestands. Dennoch überwiegen meist die Vorteile, die das Indizieren mit sich bringt.

Insbesondere die erste Forderung nach Vermeidung von Redundanzen führt dazu, bei der Planung die zu erhebenden Daten in mehrere Tabellen aufzuteilen. Dazu später mehr.

### Aufbau einer relationalen Datenbank

In einer relationalen Datenbank werden die Daten nicht hierarchisch in einer Datei gespeichert, sondern nach Themenkreisen (**Entitäten**) in Form von Tabellen abgelegt.

Personaldaten

PNr	Name
121	Eifrig
134	Schmidt
155	Huber

Zeitschriftendaten

ZNr	Titel
100	Umweltschutz
102	Arbeitstechnik
105	Management

Entleihdaten

PNr.	ZNr	Datum
121	100	10.01.97
121	102	10.03.97
121	102	01.02.97
134	100	03.04.97
155	105	04.04.97
155	105	01.08.97



### Aufgaben einer Datenbank

- Sie soll dem Benutzer den Zugriff auf die gespeicherten Daten ermöglichen, ohne dass dieser wissen muss, wie die Daten im System organisiert sind.

- Sie muss verhindern, dass ein Benutzer Daten sichten oder manipulieren kann, für die er keine Zugriffsberechtigung hat. Außerdem darf es nicht passieren, dass wegen Fehlmanipulationen des Benutzers Daten zerstört werden oder gar der ganze Datenbestand unbrauchbar wird.
- Es muss möglich sein, die interne Datenorganisation ändern zu können, ohne dass der Benutzer seine Anwenderprogramme (Applikationen) anpassen muss. Im Idealfall merkt der Benutzer von der Strukturänderung nichts.

### Anforderungen an ein Datenbankmodell

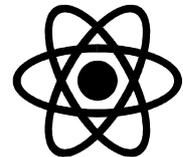
Es muss **Datenintegrität** herrschen, d.h. alle Informationen in einer Datenbank müssen nach einem einheitlichen Schema gespeichert sein. Die **Datenkonsistenz** muss gewährleistet sein. Das bedeutet: die Kontrollinstanzen des Datenbanksystems müssen die permanente Prüfung der Daten auf Korrektheit, Vollständigkeit und Redundanz gewährleisten. Durch strikte Trennung von Datenbasis und Anwenderprogrammen muss die **Datenunabhängigkeit** sichergestellt sein, z.B. darf eine Änderung der Dateistruktur keine Änderungen in einem Anwenderprogramm nach sich ziehen. Da der Zugriff auf die Daten durch das Datenbankverwaltungssystem erfolgt, kann hiermit die **Einhaltung der Datenschutzvorschriften** überwacht werden. Die Zugriffsrechte werden unabhängig von Daten und Anwendungsprogrammen vergeben. Jede Datenbank hat ihre eigene **Datenbanksprache** (Kommandosprache). Eine häufig verwendete Sprache ist SQL (Structured Query Language).

### Objekte einer Datenbank

Objekte haben oft sehr viele Merkmale. In eine Datei sollte man nur die Merkmale aufnehmen, die für den jeweiligen Zweck der Datei auch benötigt werden. Dabei darf **jedes Merkmal nur eine einzige Eigenschaft** des Objekts erfassen. Beispiel: Die Eigenschaft 'Name' einer Person wäre ungünstig, da der Name aus Vor- und Nachnamen besteht; d.h. hieraus sollten zwei Merkmale gemacht werden.

Man spricht von **Atomisierung von Merkmalen**.

Die Merkmale können eine **Beziehung zueinander** haben; z.B. können sie alle möglichen Eigenschaften einer bestimmten Person beschreiben. Solche Merkmale gehören zusammen und sollten deshalb in der Datei auch als zusammengehörig erkannt werden. Man fasst sie deshalb zu einem **Datensatz** zusammen.



Der Computer versteht natürlich nicht, was in einem Datensatz eingetragen ist. Um die zugehörigen Informationen bereitzustellen, müssen die Datensätze und deren Merkmale (Datenfelder) mit einer Vorgabe verglichen werden. Das ist jedoch ein Problem, wenn die Vorgabe '**Hans Meier**' lautet und mehrere Personen mit diesem Namen in der Datei existieren. Das Suchergebnis wäre in einem solchen Fall **nicht eindeutig**. Deshalb gilt:

**Jeder Datensatz benötigt ein eindeutiges Merkmal!**

Leider besitzen viele Objekte keine brauchbaren Merkmale, die sie von anderen unterscheiden. Aus diesem Grund ordnet man jedem Datensatz ein **künstliches Merkmal** zu, das eindeutig ist: in der Regel ist das eine **natürliche Zahl**.

*(So wird es beispielsweise auch auf dem Personalausweis gemacht: damit man die vielen 'Hans Meier', die es in Deutschland gibt, unterscheiden zu können, hat jeder eine andere Identifikationsnummer.)*



Ein Merkmal, mit dem man die **Datensätze eindeutig unterscheiden** kann, heißt **Schlüssel**. Man spricht in diesem Zusammenhang von einer **Individualisierung von Datensätzen**.

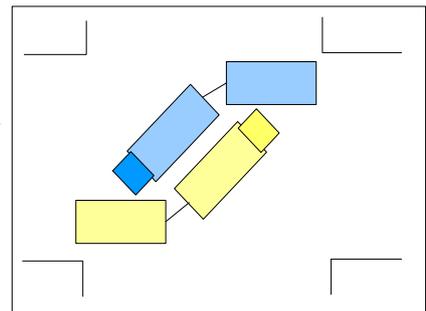
Ähnlich wie in der Mathematik jedem x-Wert ein y-Wert zugeordnet wird (Relation) wird hier jedem Schlüssel ein Datensatz zugeordnet. Dateien mit solchen Zuordnungen heißen relationale Dateien. Dabei bildet jede Datei eine Menge von Datensätzen.

### Nutzung von Datenbanken

Datenbanken werden überall dort verwendet, wo große Datenmengen zu verwalten sind, also in beinahe allen Bereichen, in denen die EDV zum Einsatz kommt. Neben den Datenbanken der öffentlichen Organe (z.B. Rentenanstalt, Arbeitsamt, Einwohnermeldeamt, Bundeskriminalamt ...) gibt es nach Sachgebieten organisierte Datenbanken, die jedermann zur Verfügung stehen, der über einen Rechner mit einem Modem verfügt. Beispiele hierfür sind juristische, medizinische, betriebswirtschaftliche und allgemein branchenspezifische Datenbanken. Für nahezu jedes Fachgebiet besteht eine Datenbank mit ständig aktuell gehaltenen Informationen.

Die **Notwendigkeit der Synchronisation** besteht bei Mehrbenutzersystemen, wenn z.B. mehrere Benutzer auf die gleichen Daten zugreifen wollen. Es müssen Vorgänge die nebeneinander laufen, zeitlich abgestimmt werden.

Die Möglichkeit **der Verklemmung** (Deadlock) kann dann auftreten, wenn zwei Benutzer  $B_1$  und  $B_2$  einer Datenbank wechselseitig auf die Datensätze  $D_1$  und  $D_2$  zugreifen wollen.  $B_1$  greift zunächst auf  $D_1$  zu und sperrt damit diesen Datensatz.  $B_2$  greift auf  $D_2$  zu und sperrt diesen Datensatz ebenfalls. Falls beide Benutzer zu diesem Zeitpunkt auf die jeweils vom anderen Benutzer gesperrten Datensätze zugreifen, blockieren sich die Prozesse gegenseitig und das System wird lahmgelegt. Auch im täglichen Leben gibt es Situationen, in denen sich zwei Abläufe gegenseitig blockieren, wie z.B. die in der Abbildung dargestellten LKWs in einer Straßenkreuzung.



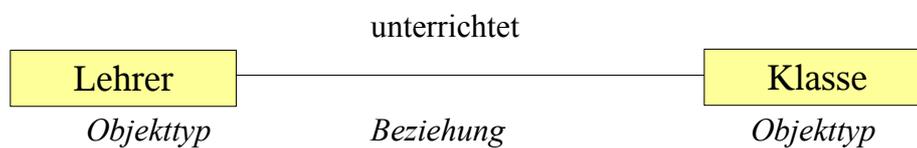
### Datenstruktur modellieren

Um eine Datenbank aufzubauen, muss ein möglichst exaktes Abbild der relevanten Welt definiert werden, ein **Datenmodell**. Ein **Modell ist eine zweckorientiert vereinfachte und strukturgleiche Abbildung der Wirklichkeit**. Zum Zweck der Verarbeitung gebildete Informationen (zweckorientiertes Wissen) heißen Daten. Der Teilausschnitt der Welt, der zu analysieren ist, wird Miniwelt genannt. Das Datenmodell soll die Miniwelt mit ihrem vollständigen Bedeutungsgehalt (semantisch vollständig) abbilden. Dazu müssen alle relevanten Objekte und Beziehungen zwischen den Objekten in einem semantischen Datenmodell erfasst und beschrieben werden.

Betrachtet man die "Miniwelt Schule", erkennt man gewisse **Objekte** wie das Fach Englisch, einen Schüler namens Meier, einen Lehrer namens Müller, einen Schulleiter namens Direx, einen Klassenleiter namens Huber, einen Raum mit der Nummer 112. Für Objekt wird auch der Begriff **Entität** oder Instanz verwendet. Um das Beispiel nicht zu komplex zu gestalten, soll eine Beschränkung auf die Objekte **Schüler, Lehrer und Klassen** vorgenommen werden. Im objektorientierten Ansatz 'besitzt' ein Objekt neben Eigenschaften auch Methoden (Operationen). Eine Klasse definiert die Eigenschaften und Methoden gleichartiger Objekte. Zwischen diesen Objekten bestehen **Beziehungen**, die bestimmte Abläufe (Prozesse) oder Abhängigkeiten in der Miniwelt darstellen. Die Unter-

richtsbelegung stellt die Beziehung zwischen einem Lehrer und den von ihm unterrichteten Klassen dar. Um zum Datenmodell zu gelangen, stellt man Objekte und Beziehungen mit gleichartigen **Attributen** zu Typen zusammen. Somit ergeben sich für das gewählte Beispiel

- die Objekttypen **Lehrer und Klassen**
- die Beziehung **unterrichtet**



Alle Lehrer bilden die Objektmenge **Lehrer**, alle Klassen die Objektmenge **Klasse** und alle Unterrichtsbelegungen die Beziehungsmenge **unterrichtet**. Jeder Objekt und Beziehungstyp definiert bestimmte **Eigenschaften (Merkmale, Attribute)**.

- Die Attribute des Objekttyps Lehrer sind:  
*Lehrernummer, Nachname, Vorname, Amtsbezeichnung, Fächerkombination.*
- Für den Objekttyp Klasse lassen sich z. B. folgende Eigenschaften feststellen: *Klassenbezeichnung, Ausbildungsrichtung, Klassenleiter L-Nr, Klassensprecher S-Nr.*
- Dem Beziehungstyp **unterrichtet** lassen sich z. B. folgende Merkmale zuordnen:  
*L-Nr, Klassenbez, Fach*

Für eine solche Datenmodellierung gibt es eine Reihe von Techniken, die je nach den Eigenheiten des Systems Vor- und Nachteile haben. Im Datenbankbereich ist die Entity-Relationship-Modellierung die mit Abstand häufigste und erfolgreichste Technik. Die Elemente des Entity-Relationship-Modells sind **Entitäten, Beziehungen zwischen Entitäten und Attribute** zur Charakterisierung von Entitäten.

### Das Entity-Relationship-Modell (ERM)

Wichtige Begriffe in der Datenbanktheorie in alphabetischer Reihenfolge:

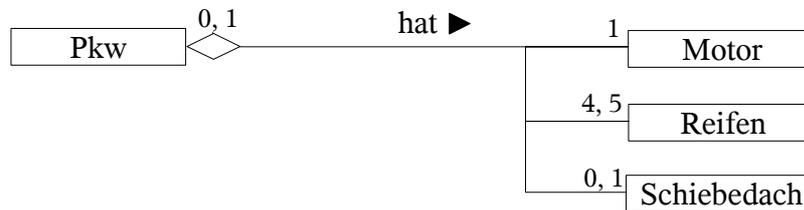
- **Attribut (Spaltenname)**: Das Attribut entspricht einem Merkmal eines Tupels und beschreibt somit eine spezifische Eigenschaft einer Entitätsmenge.
- **Attributwert (Wert, Datum)** : Datenwert, der das zugehörige Attribut eines Tupels beschreibt.
- **Domäne (Wertebereich)**: Gewisse Attribute schränken die zugehörigen Attributwerte ein (z.B. Wochentag --> Montag, ... ). Diese Wertebeschränkung nennt man Domäne oder Wertebereich.
- **Entität (Tabellenname)**: Eine Entität stellt einen Themenkreis dar, welcher Elemente mit gleichen Merkmalen umfasst. z.B. : Personen, Klassen, Kurse, Autos
- **Entitätsmenge (Datensätze)**: Die Entitätsmenge beinhaltet alle zu den Merkmalen einer Entität gehörenden Werte. Sie entspricht allen gespeicherten Datensätzen einer Tabelle.
- **Nullwerte**: Wenn ein Attribut eines Tupels einen Nullwert enthält, so bedeutet dies, dass dieses Attribut keinen Attributwert besitzt und somit keine Information beinhaltet. Die Zahl Null stellt eine Information dar, der Nullwert jedoch nicht.
- **Tabelle**: Eine Tabelle umfasst eine Entität mit der dazugehörigen Entitätsmenge. Man versteht darunter eine komplette Tabelle mit Entitätsbezeichnungen Attributen und Tupel.



In diesem Zusammenhang unterscheidet man noch folgende Sonderfälle:

**Aggregation:** Zusammensetzung von Teilen zu einem Ganzen (Spezialfall der Assoziation). Die verbindende Assoziationslinie endet bei der Aggregation mit einer Raute an der Seite des Ganzen.

Beispiel:



**Komposition:** Die Einzelteile sind vom Ganzen existenzabhängig (Spezialfall der Aggregation). Die verbindende Assoziationslinie endet bei der Aggregation mit einer gefüllten Raute an der Seite des Ganzen.



Um Beziehungen zwischen den Tabellen herstellen zu können, müssen zwei Begriffe bekannt sein:

- Identifikationsschlüssel (ID - Schlüssel)**

Jedes Tupel (Datensatz) muss eindeutig identifizierbar sein. Das kann durch **ein Attribut oder einer Kombination von Attributen** gewährleistet werden (z.B. Person in einer Firma -> Personalnummer) Jedem neuen Tupel muss sofort der entsprechende Attributwert des Identifikationsschlüssel zugeteilt werden können. Der Identifikationsschlüssel eines Tupels darf während dessen Existenz nicht ändern. Handelt es sich beim "Identifikationsschlüssel" um ein Attribut, so spricht man von **Primärschlüssel**, bei mehreren Attributen von **Kombinationsschlüssel**.

- Fremdschlüssel**

Ein Fremdschlüssel in einer Tabelle T2 ist ein Attribut, welches in einer Tabelle T1 den Identifikationsschlüssel bildet. Ein Fremdschlüssel kann nur diejenigen Attributwerte annehmen, welche bereits im Identifikationsschlüssel der Tabelle T1 existieren.

Kurzschreibweise für Tabellen: **Entitätsname**(ID - Schlüssel, Attribut1, Attribut2, Attribut3, ...)

Mögliche Beziehungstypen:

	<b>1</b>	<b>c</b>	<b>m</b>	<b>mc</b>	Beziehungen:
<b>1</b>	1 - 1	c - 1	m - 1	mc - 1	← <b>hierarchisch</b>
<b>c</b>	1 - c	c - c	m - c	mc - c	← <b>konditionell</b>
<b>m</b>	1 - m	c - m	m - m	mc - m	← <b>netzwerkartig</b>
<b>mc</b>	1 - mc	c - mc	m - mc	mc - mc	

**Im ERM sind nur hierarchische Beziehungen erlaubt!**

## Die 1 - 1 Beziehung



Jede Person besitzt **genau ein** Auto bzw. jedes Auto **gehört genau** einer Person.

### Personen

PNr	Name	Vorname	ANr
1	Eitel	Emil	3
2	Immergrün	Ingrid	5
3	Leicht	Lolita	2
4	Macho	Manuel	1
5	Prächtig	Paula	4

↑ ID - Schlüssel      Fremdschlüssel      ↑

### Autos

ANr	Automarke	Typ	Baujahr
1	Audi	Quattro	1994
2	Opel	Corsa	1996
3	Porsche	Carrera	1995
4	Volkswagen	Golf	1993
5	Fiat	Uno	1992

## Die 1 - c Beziehung



Jede Person besitzt **kein oder genau ein** Auto bzw. jedes Auto **gehört genau** einer Person.

### Personen

PNr	Name	Vorname	ANr
1	Eitel	Emil	3
2	Immergrün	Ingrid	5
3	Leicht	Lolita	2
4	Macho	Manuel	1
5	Prächtig	Paula	4

### Autos

ANr	Automarke	Typ	Baujahr	PNr
1	Audi	Quattro	1994	4
2	Opel	Corsa	1996	3
3	Porsche	Carrera	1995	1

**Nullwerte in Fremdschlüsseln sind im ER - Datenmodell unzulässig.**

## Die 1 - m Beziehung



Jede Person besitzt **mindestens ein** Auto bzw. jedes Auto **gehört genau** einer Person.

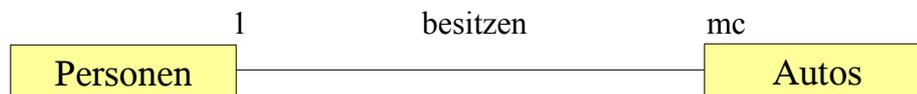
**Personen**

PNr	Name	Vorname	ANr
1	Eitel	Emil	3
2	Immergrün	Ingrid	5
3	Leicht	Lolita	2
4	Macho	Manuel	1
5	Prächtigt	Paula	4

**Autos**

ANr	Automarke	Typ	Baujahr	PNr
1	Audi	Quattro	1994	4
2	Opel	Corsa	1996	3
3	Porsche	Carrera	1995	1
4	Volkswagen	Golf	1993	5
5	Fiat	Uno	1992	2
6	Lancia	Thema	1990	1
7	BMW	M3	1995	4

Die Tabelle Autos muss mindestens gleich viele Tupel besitzen wie die Tabelle Personen. Der Fremdschlüssel PNr in der Tabelle Autos kann den gleichen Attributwert mehrfach annehmen.

**Die 1 - mc Beziehung**

Jede Person kann **beliebig viele** Autos besitzen (0, 1 oder mehr) bzw. jedes Auto **gehört genau** einer Person.

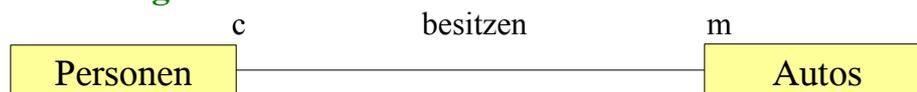
**Personen**

PNr	Name	Vorname	ANr
1	Eitel	Emil	3
2	Immergrün	Ingrid	5
3	Leicht	Lolita	2
4	Macho	Manuel	1
5	Prächtigt	Paula	4

**Autos**

ANr	Automarke	Typ	Baujahr	PNr
1	Audi	Quattro	1994	4
2	Opel	Corsa	1996	3
3	Porsche	Carrera	1995	1
4	Volkswagen	Golf	1993	5
5	Jeep	Cherokee	1997	3
6	Lancia	Thema	1990	1
7	BMW	M3	1995	4

In der Tabelle Personen könne nun auch Tupel existieren, deren ID-Schlüsselwert nicht im Fremdschlüssel PNr in der Tabelle vorkommt.

**Die c - m Beziehung**

Jede Person besitzt **mindestens ein** Auto bzw. jedes Auto hat entweder **keinen oder genau einen** Besitzer.

**Personen**

PNr	Name	Vorname	ANr
1	Eitel	Emil	3
2	Immergrün	Ingrid	5
3	Leicht	Lolita	2

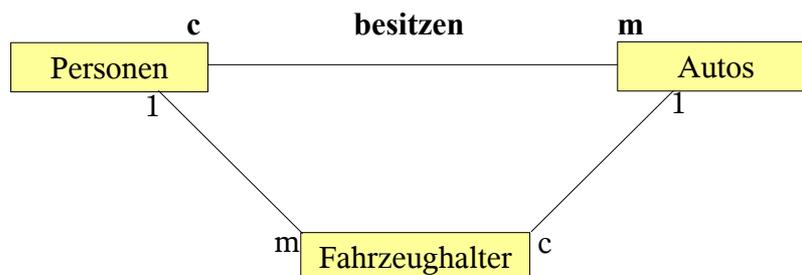
**Autos**

ANr	Automarke	Typ	Baujahr	PNr
1	Audi	Quattro	1994	3
2	Opel	Corsa	1996	2
3	Porsche	Carrera	1995	1
4	Volkswagen	Golf	1993	
5	Fiat	Uno	1992	
6	Lancia	Thema	1990	1
7	BMW	M3	1995	3

Bei dem **Fremdschlüssel PNr** in der Tabelle Autos ergeben sich bei dieser Darstellung **Nullwerte**. Bei der Variante, bei der Personen mit mehreren Autos in der Tabelle Personen mehrfach eingetragen sind, würde dies zu **nicht mehr eindeutigen ID-Schlüsseln** führen.

Abhilfe:

Die konditionelle c - m Beziehung wird in zwei hierarchische Beziehungen, 1 - m bzw. 1 - c transformiert (siehe nachfolgende Abbildung).



Jede Person ist **mindestens ein** Fahrzeughalter Auto und somit besitzt jede Person **mindestens ein** Auto. Jedes Auto hat entweder **keinen oder genau einen** Fahrzeughalter, und damit einen Besitzer.

Die drei Tabellen haben nun folgenden Aufbau:

**Personen**

PNr	Name	Vorname	ANr
1	Eitel	Emil	3
2	Immergrün	Ingrid	5
3	Leicht	Lolita	2

**Autos**

ANr	Automarke	Typ	Baujahr
1	Audi	Quattro	1994
2	Opel	Corsa	1996
3	Porsche	Carrera	1995
4	Volkswagen	Golf	1993
5	Fiat	Uno	1992
6	Lancia	Thema	1990
7	BMW	M3	1995

Kurzschreibweise:

**Personen** (PNr, Name, Vorname)

**Autos** (ANr, Marke, Typ)

**Fahrzeughalter** (ANr, PNr)

Fahrzeughalter	
PNr	ANr
1	3
1	6
2	2
3	1
3	7

### Der Normalisierungsprozess

Die Normalisierung bezweckt die redundanzfreie Speicherung von Informationen innerhalb der Tabellen der Datenbasis.

**Redundanzfreie Datenspeicherung** bedeutet, dass kein Teil eines Datenbestandes weggelassen werden kann, ohne dass dies zu Informationsverlusten führt. Sie bringt Speicherplatzersparnis.

Der Normalisierungsprozess führt in der Regel zu Aufteilung der Daten in verschiedene Tabellen.

Im folgenden **Beispiel** werden für jeden Mitarbeiter einer Firma dessen Abteilung und seine Projektdaten erfasst.

Name	Abteilung	Projekt-Nr.	Beschreibung	Stunden	Tätigkeit	Stundenlohn
Richter, Hans	Personal	1	Verkaufspromotion	63	Koordination	15,-- €
Metzger, Helga	Einkauf	2	Konkurrenzanalyse	32	Erstellung	14,-- €
Schneider, Udo	Einkauf	2, 3	Konkurrenzanalyse, Kundenumfrage	44, 51	Erstellung, Durchführung	14,-- €, 18,-- €
Wild, Inge	Personal	1, 2	Verkaufspromotion, Konkurrenzanalyse	0, 108	Betreuung, Erstellung	12,-- €, 14,-- €
Förster, Irene	Buchhaltung	1, 2, 3	Verkaufspromotion, Konkurrenzanalyse, Kundenumfrage	28, 40, 98	Durchführung, Erstellung, Vorbereitung	18,-- €, 14,-- €, 13,-- €
Lang, Felix	Verkauf	1, 3	Verkaufspromotion, Kundenumfrage	87, 93	Durchführung, Durchführung	18,-- €, 18,-- €

Diese Tabelle bezeichnet man als **nicht-normalisiert!** Es sind in mehreren Feldern eines Datensatzes mehrere Informationen gespeichert worden. Das führt gewöhnlich zu Problemen bei den Datenbankoperationen **Ändern, Einfügen oder Löschen** von Datensätzen, z.B. gehen auch Projektdaten verloren, wenn ein Mitarbeiter aus der Firma ausscheidet und sein Datensatz gelöscht wird. Bei der Änderung einer Projektbeschreibung müssen alle Datensätze daraufhin durchgesehen und abgeändert werden.

**Der Normalisierungsprozess verläuft über die Bildung sogenannter Normalformen, die aufeinander aufbauen.**

## 1. Normalform

Eine Tabelle befindet sich dann in der **1. Normalform**, wenn alle Attribute nur einfache (atomare) Attributwerte aufweisen.

M_Nr	Name	Vorname	A_Nr	Bezeichnung	P_Nr	Beschreibung	Std.	Tätigkeit	Std.-Lohn
1	Richter	Hans	1	Personal	1	Verkaufspromotion	63	Koordination	15,-- €
2	Metzger	Helga	2	Einkauf	2	Konkurrenzanalyse	32	Erstellung	14,-- €
3	Schneider	Udo	2	Einkauf	2	Konkurrenzanalyse	44	Erstellung	14,-- €
3	Schneider	Udo	2	Einkauf	3	Kundenumfrage	51	Durchführung	18,-- €
4	Wild	Inge	1	Personal	1	Verkaufspromotion		Betreuung	12,-- €
4	Wild	Inge	1	Personal	2	Konkurrenzanalyse	108	Erstellung	14,-- €
5	Förster	Irene	3	Buchhaltung	1	Verkaufspromotion	28	Durchführung	18,-- €
5	Förster	Irene	3	Buchhaltung	2	Konkurrenzanalyse	40	Erstellung	14,-- €
5	Förster	Irene	3	Buchhaltung	3	Kundenumfrage	98	Vorbereitung	13,-- €
6	Lang	Felix	4	Verkauf	1	Verkaufspromotion	97	Durchführung	18,-- €
6	Lang	Felix	4	Verkauf	3	Kundenumfrage	93	Durchführung	18,-- €

Nach Einfügung der Attribute M\_Nr, A\_Nr sowie P\_Nr können Mitarbeiter, Abteilung und Projekt klar identifiziert werden. Alle Attribute weisen nur noch einfache Attributwerte auf, wobei Nullwerte (kein Eintrag im Datenfeld) ebenfalls Attributwerte darstellen. Es sind nun aber Redundanzen eingetreten. Außerdem ist ersichtlich, dass es innerhalb der Tabelle verschiedene Sachgebiete gibt, die unabhängig voneinander existieren können.

## 2. Normalform

Eine **Tabelle** befindet sich dann in der **2. Normalform**, wenn sie schon in der **1. Normalform** ist und jedes nicht zum Primärschlüssel gehörende Attribut voll vom Primärschlüssel abhängig ist.

Mitarbeiter

M_Nr	Name	Vorname
1	Richter	Hans
2	Metzger	Helga
3	Schneider	Udo
4	Wild	Inge
5	Förster	Irene
6	Lang	Felix

Abteilungen

A_Nr	Bezeichnung
1	Personal
2	Einkauf
3	Buchhaltung
4	Verkauf

## Projekte

P_Nr	Beschreibung	Std.	Tätigkeit	Stundenlohn	M_Nr
1	Verkaufspromotion	63	Koordination	15,-- €	1
2	Konkurrenzanalyse	32	Erstellung	14,-- €	2
2	Konkurrenzanalyse	44	Erstellung	14,-- €	3
3	Kundenumfrage	51	Durchführung	18,-- €	3
1	Verkaufspromotion		Betreuung	12,-- €	4
2	Konkurrenzanalyse	108	Erstellung	14,-- €	4
1	Verkaufspromotion	28	Durchführung	18,-- €	5
2	Konkurrenzanalyse	40	Erstellung	14,-- €	5
3	Kundenumfrage	98	Vorbereitung	13,-- €	5
1	Verkaufspromotion	97	Durchführung	18,-- €	6
3	Kundenumfrage	93	Durchführung	18,-- €	6

Kurzschreibweise:

**Mitarbeiter** (M\_Nr, Name, Vorname)

**Abteilungen** (A\_Nr, Bezeichnung)

**Projekte** (P\_Nr, Beschreibung, Stunden, Tätigkeit, Stundenlohn, M\_Nr)

Gemäß der 2. Normalform sind nur Tabellen maßgebend, die als Primärschlüssel eine Attributkombination enthalten. Die Tabellen Mitarbeiter und Abteilungen entsprechen diesen Erfordernissen. Die Attribute sind jeweils voll von den Primärschlüsseln M\_Nr bzw. A\_Nr abhängig. In der Tabelle Projekte ist aber der Stundenlohn von der Tätigkeit und nicht vom Primärschlüssel abhängig. Außerdem sind immer noch Redundanzen vorhanden.

### 3. Normalform

Eine **Tabelle** befindet sich dann in der **3. Normalform**, wenn sie schon in der 2. Normalform (bzw. mit einfachem Primärschlüssel in der 1. Normalform) ist und kein Nichtschlüselfeld von einem anderen Nichtschlüselfeld abhängig ist.

Diese Definition bedeutet, dass die Attribute innerhalb einer Tabelle nur vom ID-Schlüssel funktional abhängig sind und untereinander keine sonstigen funktionalen Abhängigkeiten existieren. Da in der Tabelle Projekte der Stundenlohn von der Tätigkeit abgeleitet werden kann, muss diese Tabelle weiter aufgeteilt werden.

Projekte

P_Nr	Beschreibung
1	Verkaufspromotion
2	Konkurrenzanalyse
3	Kundenumfrage

Projektauswertung

P_Nr	M_Nr	Stunden	T_Nr
1	1	63	1
2	2	32	2
2	3	44	2
3	3	51	4
1	4	56	5

Tätigkeiten

T_Nr	Tätigkeit	Std.-Lohn
1	Koordination	15,-- €
2	Erstellung	14,-- €
3	Vorbereitung	13,-- €
4	Durchführung	18,-- €
5	Betreuung	12,-- €

2	4	108	2
1	5	28	4
2	5	40	2
3	5	98	3
1	6	97	4
3	6	93	4

Kurzschreibweise:

Projekte (P\_Nr, Beschreibung)

Tätigkeiten (T\_Nr, Tätigkeit, Stundenlohn)

Projektauswertung (P\_Nr, M\_Nr, Stunden, T\_Nr)

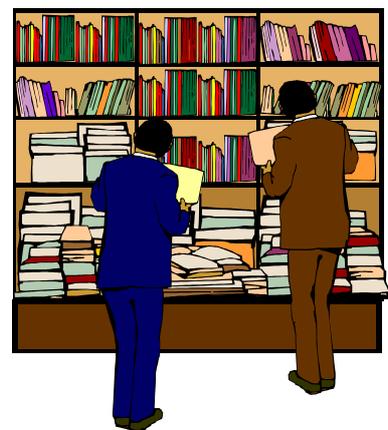
Tabellen, die sich in der 3. Normalform befinden, werden als normalisiert bezeichnet. Die darin enthaltenen Informationen sind weitgehend redundanzfrei.

**Aufbau einer Datenbank für eine Bücherei**

Dieses Beispiel soll den Aufbau einer Datenbank zeigen und zwar von den Vorüberlegungen bis hin zur Umsetzung am Computer.

Für eine Bücherei sollen folgende Daten erfasst werden:

- x Buchtitel
- x Buchautor
- x Erscheinungsjahr
- x Verlag
- x Genre
- x Lesername
- x Leseranschrift
- x Leserstatus (Normal, Schüler/Student, Rentner)
- x Jahresgebühr (15,-- €, 0,-- €, 5,-- €)
- x Entleihdatum
- x Rückgabedatum



Die Daten könnten somit folgendermaßen in einer Tabelle gespeichert werden:

Müller	Schweinfurt	normal	15,- €	Die Legende vom Boomerang Hobbylexikon Modelleisenbahn Schülerduden Physik	Adler Albrecht Ahlheim	1966 1976 1989	Holz rororo Dudenverl.	Kinder Technik Physik.	03.05.02 03.05.02 02.05.02	22.05.02 22.05.02
Schmidt	Gochsheim	Schüler	0,- €	Vorstadtkrokodile One hundred easy stories	Grün Müller	1978 1959	rororo Hueber	Kinder Englisch	05.05.02 05.05.02	11.05.02
Koch	Niederwerrn	Rentner	5,- €	Spiegel der Zeiten Die Magermilch-Bande	Mager Baer	1983 1981	Diesterweg Fischer	Geschichte Belletristik	06.05.02 06.05.02	
Müller	Üchtelhausen	normal	15,- €	Von den Steinzeitjägern	Beyerlein	1991	Oetinger	Kinder	06.05.02	09.05.02
Meier	Schweinfurt	normal	15,- €	Ein Kapitel für sich	Kempowski	1980	dtv	Belletristik	07.05.02	25.05.02
Schmidt	Gochsheim	Schüler	0,- €	Keine Hosenträger für Opa	Banscherus	1985	Arena	Jugend	11.05.02	

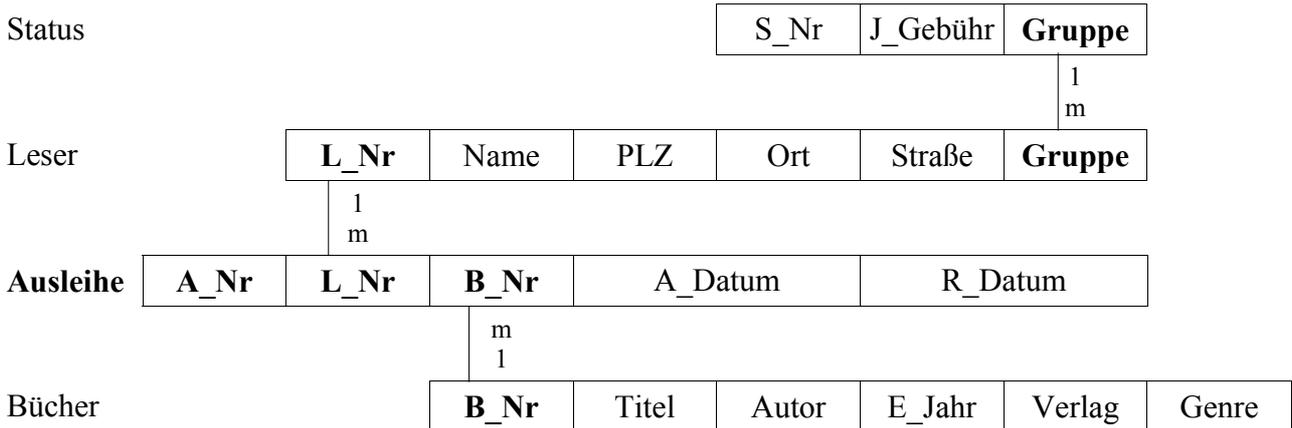
Man erkennt, dass mit dieser Art der Speicherung eine Reihe von Problemen entstehen würden:

- Es ist noch kein Primärschlüssel vorhanden.
- Leserdaten werden mehrfach gespeichert. (Redundanz)
- Es können sich leicht Schreibfehler einschleichen, insbesondere bei Namen wie Meier, Maier, Meyer ... (Integritätsproblem)
- Ändert sich die Anschrift eines Lesers, so muss diese in allen Datensätzen korrekt ausgebessert werden. Kein Datensatz darf vergessen werden. (Integritätsproblem)
- Neue Leser bzw. neue Bücher können nur mit Entleihvorgängen gespeichert werden.
- Bei jedem Entleihvorgang müssen die gesamten Daten eines Lesers bzw. eines Buches wieder vollständig mit aufgenommen werden.
- Will man einen Leser löschen, so würden evtl. auch Daten von Büchern mit gelöscht werden. Entsprechend verhält es sich, wenn man ein Buch löschen will.

Das erste Problem lässt sich leicht lösen: Wir legen für jeden Entleihvorgang eine laufende Nummer fest. Um die anderen Probleme zu beseitigen, müssen wir die Normalisierungsregeln anwenden:

1. Normalisierung: Die Datenfelder, die in einem Datensatz Mehrfachbelegungen aufweisen, müssen in eine separate Datei ausgelagert werden. In unserem Fall sind das zunächst die Bücher; d.h. wir legen für die Bücher eine neue Tabelle an. Wir erhalten dadurch zwei Tabellen:  
**Ausleihe(A\_Nr, L\_Name, L\_Anschrift, L\_Status, Gebühr, B\_Nr, A\_Datum, R\_Datum)**  
**Bücher(B\_Nr, Titel, Autor, E\_Jahr, Verlag, Genre)**
2. Normalisierung: Nur das Entleih- und das Rückgabedatum hängen vom Primärschlüssel ab. Das bedeutet, dass die Entleihdaten ebenfalls in einer separaten Datei gespeichert werden. In unserem Fall entstehen somit 3 Tabellen:  
**Ausleihe(A\_Nr, L\_Nr, B\_Nr, A\_Datum, R\_Datum)**  
**Bücher(B\_Nr, Titel, Autor, E\_Jahr, Verlag, Genre)**  
**Leser(L\_Nr, Name, Anschrift, Status, Gebühr)**
3. Normalisierung: Das Datenfeld Gebühr hängt nicht vom Schlüsselfeld 'L\_Nr', sondern nur vom Nichtschlüsselfeld 'Status' ab. Deshalb wird auch hierfür eine weitere Tabelle erstellt.

Nach Anwendung aller drei Normalisierungsregeln ergeben sich folgende Tabellen mit den entsprechenden Verknüpfungen:



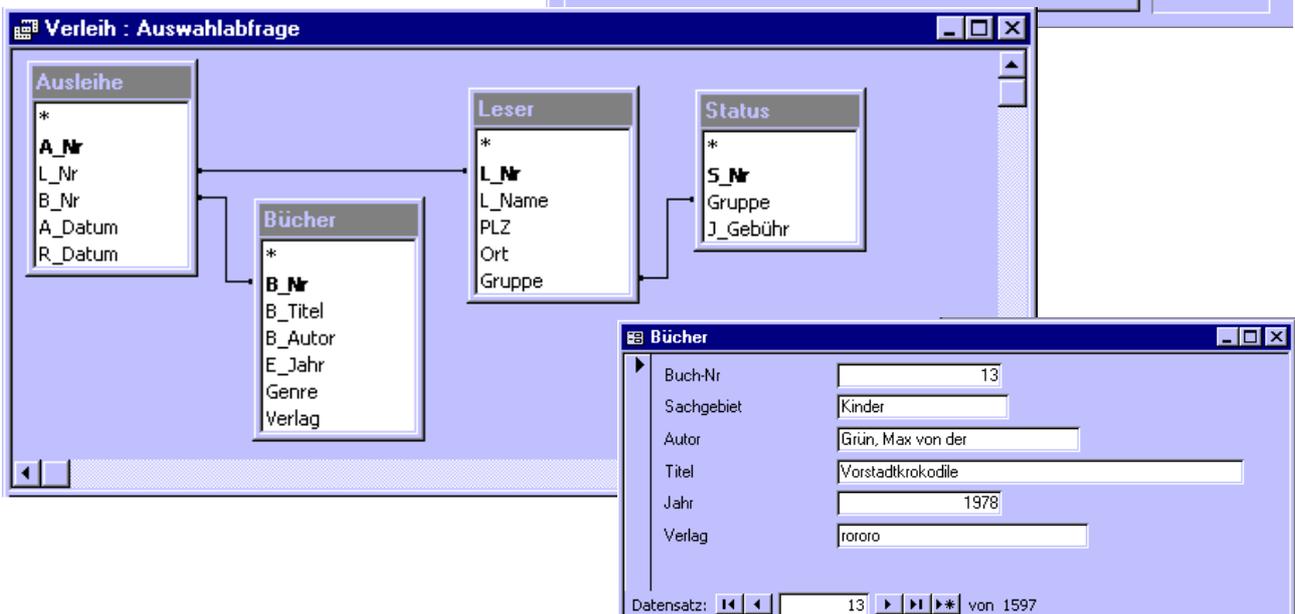
Wichtig sind die gemeinsamen Felder, über die die Tabellen miteinander verknüpft werden können. In der Haupttabelle 'Ausleihe' ist nur das Feld 'A\_Nr' Primärschlüssel, die Felder L\_Nr und B\_Nr sind Schlüsselfelder in anderen Tabellen. In diesem Fall spricht man von **Fremdschlüssel**.

Die Abbildung zeigt den Tabellenentwurf der Tabelle 'Bücher' in der Datenbankanwendung Access.

Nach dem Anlegen der einzelnen Tabellen werden sie noch mit Inhalt gefüllt, d.h. es können sofort (oder auch später) entsprechende Datensätze angelegt werden.

Die Verknüpfung der Tabellen erhält man mit Hilfe einer Abfrage.

Eine Abfrage ermöglicht auch **Selektion** und **Projektion** von Datensätzen bzw. Datenfeldern. Weitere Möglichkeiten bieten die Datenbankanwendungen durch die Objekte Formular (z.B. Bildschirmmasken), Bericht (Ausgabe über Drucker - z.B. Leserausweis) usw.



Programmintern wird die Datenbankabfragesprache SQL verwendet, auch wenn z.B. eine Abfrage mittels Menü erzeugt wird. Obige Abfrage lautet in SQL:

```
SELECT FROM ((Ausleihe INNER JOIN Bücher ON Ausleihe.B_Nr = Bücher.B_Nr) INNER JOIN Leser ON Ausleihe.L_Nr = Leser.L_Nr) INNER JOIN Status ON Leser.Gruppe = Status.Gruppe;
```

Die Sprachelemente von SQL lassen sich in zwei Kategorien einteilen: Sprachelemente für die Definition und für die Manipulation der Daten. Dabei werden die Daten durch logische Bedingungen charakterisiert, der Weg zu den Daten wird nicht angegeben. SQL verarbeitet immer eine Menge von Datensätzen; d.h. das Ergebnis einer Abfrage ist immer eine Tabelle mit keiner, einer oder mehreren Zeilen.

## NETZWERKE

Man unterscheidet bezüglich der Ausdehnung bzw. Ausbreitung der Netze vier Arten:

**LAN** - Das Lokal Area Network (Nahbereichsnetz) ist meist auf ein Haus oder ein Grundstück begrenzt.

**MAN** - Das Metropolitan Area Network (Stadtnetz) findet seinen Einsatz in Stadtgebieten.

**WAN** - Das Wide Area Network (Weitverkehrsnetz) verbindet die Computer eines ganzen Staates.

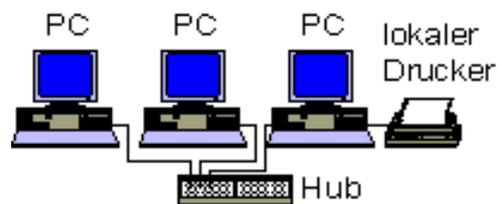
**GAN** - Das Global Area Network ist eine Spezialform des WAN und benutzt zur Datenübertragung Satellitenverbindungen und Unterseekabel.

### Netzwerkarchitekturen

Auf welche Art die einzelnen Teilnehmer eines Netzes verwaltet werden, beschreibt die Architektur des Netzwerkes.

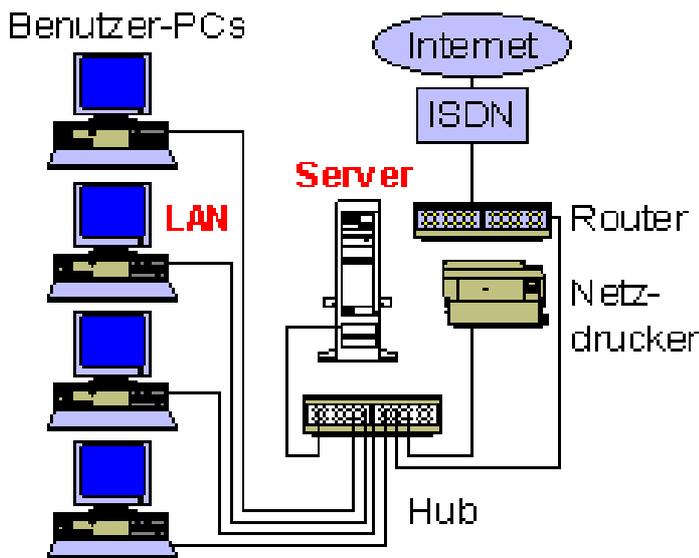
#### **Peer to Peer**

Bei dieser Netzwerkart ist jeder Rechner gleichberechtigt. Durch Zuweisung von Laufwerksbuchstaben kann jeder Computer die Daten und die Speichermedien (CD-ROM, Festplatten) des anderen nutzen. Ein bekannter Vertreter ist Windows 98 oder Windows ME.



#### **Client-Server-Architektur**

Jedes Netzwerk besteht aus mehreren Rechnern, die gemeinsame Dienste nutzen. Allgemein kann man den Rechner, der diesen Dienst zur Verfügung stellt, als **Server** und denjenigen, der ihn nutzt, als **Client** bezeichnen. Dieser Rechner stellt dem gesamten Netz seine Dienste zur Verfügung. Da ein solcher Server i. d. R. im Dauerbetrieb läuft, ist eine Nutzung jederzeit möglich. Zum Aufbau eines solchen Servers ist ein spezielles Betriebssystem notwendig. Im Schulbereich sind Novell Netware, Windows-NT bzw. 2000 und Linux üblich. Bei der Planung von Netzwerken muss dabei auf die Betriebssysteme der Clients geachtet werden, denn das Netzwerkbetriebssystem muss das System der Clients unterstützen. Soweit auf den Clients ein Windows-Betriebssystem läuft, sind alle genannten Serversysteme möglich. Damit der Arbeitsplatzrechner den Server versteht, ist bei manchen Netzwerksystemen auch auf dem Client eine spezielle Software notwendig.



Aufgaben des Netzwerkbetriebssystems sind die Verwaltung der Daten aller Nutzer in entsprechenden Verzeichnissen und insbesondere die Vergabe von Rechten an die Benutzer. Zu unterscheiden ist nach Lese-, Schreib- und Ausführungsrechten für Dateien und für Verzeichnisse, oft auch für einzelne Maschinen (Computer). Nur über diese Rechte lässt sich ein Netzwerk sicher gestalten.

Wichtigste Aufgabe des Servers ist es, gemeinsame Dateien und Druckdienste zu verwalten. Dementsprechend muss die Struktur und der Aufbau des Servers gestaltet sein. Die wichtigsten Systeme sind heute **Fileserver-Systeme** und **Client-**

### Server-Systeme.

Der Name **Fileserver** gibt schon einen Hinweis auf die Struktur, die dahinter steckt. Der Fileserver speichert alle Daten, die allgemein von Interesse sind. Auf den Festplatten des Rechners liegt etwa die Kundendatei. Der Fileserver arbeitet nun wie eine entfernte Festplatte. Anforderungen, bestimmte Dateien zu lesen oder zu schreiben, leitet der Arbeitsplatzrechner über das Netz an den Fileserver weiter. Dieser schiebt die gewünschten Dateien samt und sonders über das Netz.

Aufgrund der geringen Datenübertragungsrate des Netzwerkes wirkt sich der Flaschenhals „Netzwerk“ noch gravierender aus als etwa der Engpass „Bussystem“ innerhalb des PC. Noch schlimmer wird die Geschichte, wenn große Datenbanken auf dem Fileserver lagern. Die Abfrage einer kurzen Zeichenfolge hat dann die Konsequenz, dass die gesamte Datenbank über das Netzwerk gesaugt wird.

Gerade diesen Problemen setzt die **Client-Server-Architektur** ein Ende. Spezielle Software verwaltet den Zugriff auf die Informationen, die auf dem Server lagern: ein sogenanntes Serverprogramm. Datenspeicher und Verwaltungssoftware trennt somit nicht ein Netz, sondern nur das lokale Bussystem des Servers.

Fordert ein Arbeitsplatzrechner eine bestimmte Information an, schickt er lediglich die Anforderung über das Netz. Dabei bedient sich die Software auf der Client-Seite einer speziellen Sprache, die die Serversoftware versteht. Häufig ist dies SQL (Structured Query Language). Die Serversoftware interpretiert die Anforderung, sucht die gewünschten Informationen und sendet nur diese gefilterten Daten zum Client zurück. Die Netzbelastung sinkt. Zudem ist diese Serversoftware hochspezialisiert und optimiert auch den Zugriff, etwa wenn mehrere Anfragen von verschiedenen Stationen gleichzeitig eingehen.

Als weitere Nutzungsmöglichkeiten für Server seien hier als Beispiele genannt der **Druckserver** (stellt im Netz einen oder mehrere Drucker zur Verfügung; oft auch nur als Software auf dem Server), **Proxy-Server** (speichert einmal aufgerufene Internetseiten über einen bestimmten Zeitraum und ermöglicht damit schnelleren Zugriff) und Webserver.

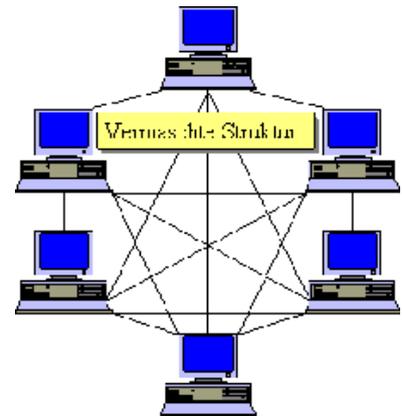
Noch ein kleiner Ausblick auf die Welt der **Großrechner**. Hier verwaltet ein Rechner mit mehreren Prozessoren nicht nur die Daten auf den Massenspeichern, sondern steuert gleichzeitig die angeschlossenen Terminals. Das Netz trennt also den Rechner von Bildschirm und Tastatur. Nur die Daten der gedrückten Tasten und die Bildschirmausgabe fließen über das Netz. Hier ist der begrenzende Faktor nicht das Netz, sondern die Zahl der Terminals, die der Rechner verwalten und versorgen muss.

## Topologien

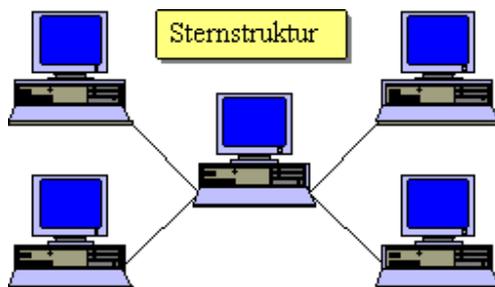
Für die Verbindungsarten bei der Vernetzung von Computern unterscheidet man vier Strukturen:

- **Vermaschte Struktur**

Jeder Teilnehmer ist mit jedem verbunden. Dieses System ist zwar sehr leistungsfähig und störungssicher, jedoch auch kostenintensiv und bedarf eines hohen Kabelaufwands.



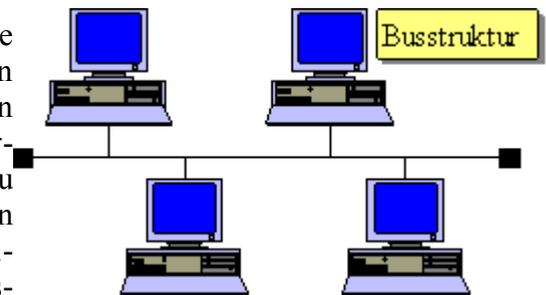
- **Sternstruktur**



Eine Zentrale versorgt alle Teilnehmer. Ein Beispiel hierfür ist das Telefonnetz der Telekom. Es ist ein kostengünstiges System, dessen Zuverlässigkeit aber von der Leistung und der Stabilität der Zentrale abhängt.

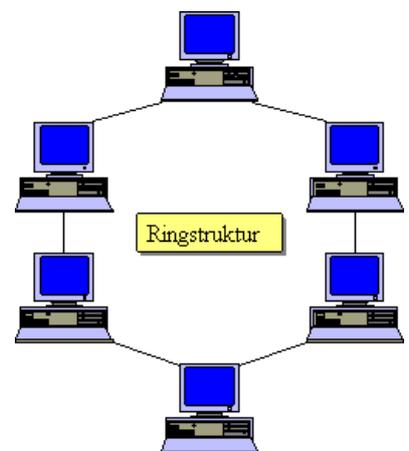
- **Busstruktur**

Bei der am meisten eingesetzten Struktur hängen alle Teilnehmer an einem passiven Kabel, das an beiden Enden mit einem Abschlusswiderstand abgeschlossen ist. Da die Daten in beide Richtungen übertragen werden, könnte es ohne den Abschluss des Kabels zu Rückkoppelungen kommen. Vorteile sind: Fällt ein Teilnehmer aus, so läuft das Netz weiter; weitere Teilnehmer können mit angeschlossen werden; es ist kostengünstig.



- **Ringstruktur (Token Ring)**

Jeder Teilnehmer hat einen Vorgänger und einen Nachfolger. Die Datenübertragung erfolgt in einer Richtung. Hierbei empfängt jeder Teilnehmer die Daten und überprüft, ob sie für ihn bestimmt sind. Ist das nicht der Fall, so verstärkt er das Signal und leitet es weiter. Ein Nachteil dieses Systems ist jedoch der Gesamtausfall des Netzes bei einer Störung eines einzelnen Teilnehmers.



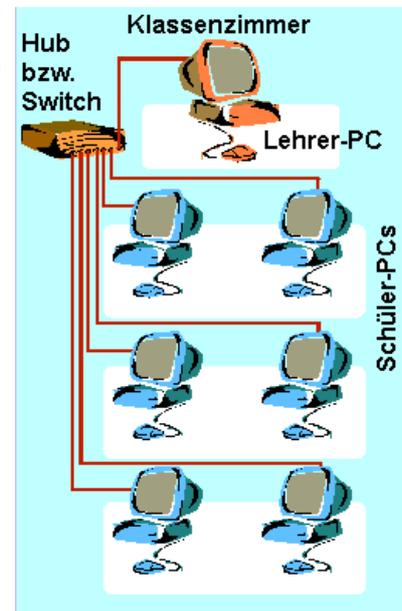
## Die Vernetzung

Bei den Techniken muss unterschieden werden, um es sich um ein LAN (Local Area Network) oder ein WAN (Wide Area Network) handelt, bzw. ob die Übertragung über Kabel oder kabellos (wireless) erfolgt.

Im Schulnetzwerk sind die Rechner meist über Hubs und Switches miteinander verbunden. Diese Geräte stellen zum einen die physikalische Verbindung her, zum anderen sorgen sie dafür, dass der Netzbetrieb mit möglichst hoher Geschwindigkeit und möglichst wenig Kollisionen ablaufen kann.

Nachdem die Rechner physikalisch miteinander verbunden sind, benötigen sie noch Software um miteinander Daten austauschen zu können oder auf gemeinsame Ressourcen wie Drucker, CD's, usw. zugreifen zu können.

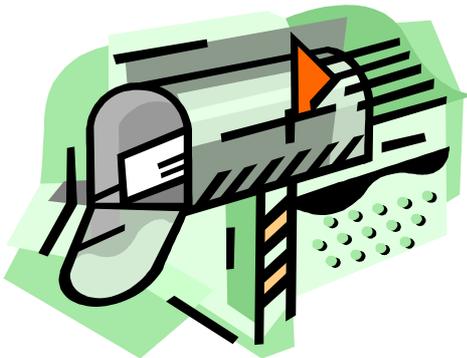
Hier sind im LAN-Bereich vor allem die **Kommunikationsprotokolle IPX/SPX (Novell) und NetBEUI (Windows)** zu nennen. **TCP/IP** ist das Kommunikationsprotokoll, wenn es um Datenaustausch über das Internet geht. Auf Grund der weiten Verbreitung des Internets entwickelt sich TCP/IP immer mehr zu einem Standard auch in LANs.



## TCP/IP-Grundlagen

### Netzwerkprotokolle

Damit Computer im Netzwerk sich miteinander verständigen können, müssen Sie die gleiche "Sprache sprechen". Es muss also verbindlich festgelegt sein, wie die Informationen von Computer A zu Rechner B kommen. Diese Aufgabe übernehmen in Netzen die Protokolle. Welches Protokoll auch immer zum Einsatz kommt, sobald ein Computer ein Teil des Netzwerks wird, erhält er eine Adresse. Genauer gesagt: Das entsprechende Netzwerkgerät, wie Ethernet- oder ISDN-Karte, Modem etc., wird mit einer Adresse versehen. Es ist ja möglich, dass ein Computer mehrere Netzwerkgeräte hat (zum Beispiel Ethernet-Karte fürs LAN und ISDN-Karte für das Internet). Somit kann ein Computer unter verschiedenen Adressen in verschiedenen Netzen bekannt sein. Über diese Adresse ist er eindeutig identifizierbar. Ähnlich wie ein Postbote einen an Sie adressierten Brief anhand Ihrer Anschrift in Ihren Briefkasten wirft, gelangen die für einen Computer im Netzwerk bestimmten Daten ebenfalls über seine Adresse an ihn. Große Netze, wie das Internet, werden in kleinere überschaubarere Einheiten aufgeteilt. Diese Einheiten sind Subnetze. Subnetze sind für sich eigenständige Netzwerke, die unabhängig funktionieren können.



### Router

Die Verbindungen zwischen diesen Subnetzen schaffen Router. Diese Router schaufeln die Daten von einem Subnetz in ein anderes. Auch hier passt die Analogie mit der guten alten Post. Wenn Subnetze Städte sind, dann sind Router die Briefzentren dieser Ortschaften. Soll ein Brief in eine andere Stadt transportiert werden, muss nicht jeder Postbote selbst in die andere Stadt pilgern. Er liefert seinen Brief beim Zentrum ab. Um weiteres muss er sich nicht kümmern. Posttransporter verfrachten die gesammelten Briefe zum Briefzentrum der passenden Stadt. Von dort aus über-

nehmen dann wieder Postboten das Ausliefern. Genau so funktionieren auch Router zwischen Subnetzen.

**Protokollschichten**

Netzwerke sind heute in Schichten (Layers) organisiert. Jede Schicht "abstrahiert" bzw. kapselt eine bestimmte Aufgabe. Der Vorteil eines solchen Vorgehens ist, dass eine Netzwerkanwendung, wie etwa der Web-Browser, nichts über die zu Grunde liegende Netzwerk-Hardware wissen muss. Dem Browser ist es einerlei, ob er über Modem, Netzwerkkarte oder Kabelmodem mit dem Internet kommuniziert. Er verwendet nur die oberste Schicht. Die Hardware, die in der untersten Schicht verwendet wird, ist ihm egal. Die Basis für solche Netzwerkarchitekturen ist das "Open Systems Interconnect Reference Model" oder kurz OSI-Referenzmodell. Dieses Architekturmodell wurde von der "International Standards Organization" (ISO) entwickelt und ist in sieben Schichten organisiert. Ganz oben liegen die Anwendungen und ganz unten die Hardware. Die Layers dazwischen erfüllen Aufgaben, wie Fehlerkorrektur, Adressierung und das Festlegen der Datenformate.

Das OSI-Referenzmodell ist für die Praxis etwas überdimensioniert. Sieben Schichten sind für eine in der realen Welt tauglichen Umsetzung des Modells gar nicht notwendig.

Anwendungs-schicht	7	Application Layer	⇒	<i>Ftp-Programm / Mail etc. (Softwareschicht)</i>
Darstellungs-schicht	6	Presentation Layer	⇒	<i>arbeitet eng mit Betriebssystem und Dateisystem zusammen und wandelt Daten in das jeweilig lesbare Format um</i>
Kommunikations-steuerung	5	Session Layer	⇒	<i>Zuständig für Einrichtung und Verwaltung einer Verbindung von zwei Computern</i>
Transportschicht	4	Transport Layer	⇒	<i>zuständig für die Reihenfolge der Pakete und Erkennung von Übertragungsfehlern</i>
Vermittlungs-schicht	3	Network Layer	⇒	<i>Versendung von Paketen an die richtige Adresse und über den richtigen Weg</i>
Sicherungs-schicht	2	Data Link Layer / Media Access Control	⇒	<i>Zerlegung der Daten in Pakete</i>
Bitübertragung	1	Physical Layer	⇒	<i>reine Hardware inkl. Kabel, Satelliten, Netzwerkkarten etc.</i>

TCP/IP-Netzwerke reduzieren daher die Anzahl der Layers auf ganze vier. Damit sind die wichtigsten Schritte für die Datenübertragung im Netzwerk einzeln gekapselt.

Da die Schichten wie gestapelt aufeinander liegen, spricht man auch vom TCP/IP-Stapel oder TCP/IP-Stack. Einen Vergleich zwischen OSI- und TCP-Schichtenmodell zeigt folgende Übersicht:

<i>OSI</i>		<i>TCP</i>
Anwendungsschicht	7	Application Layer
Darstellungsschicht	6	
Kommunikationssteuerung	5	
Transportschicht	4	Transport Layer
Vermittlungsschicht	3	Internet
Sicherungsschicht	2	Network Layer
Bitübertragung	1	

Ganz unten liegt die **Netzzugangsschicht**. Sie stellt die Verbindung zur Netzwerkhardware, wie beispielweise Modem, Ethernet- oder ISDN-Karte, her. Sie erfüllt somit die Aufgabe, die Daten

über die Hardware ins Netzwerk zu übertragen und stellt die physikalische Verbindung zum Netz her. Mit dieser Schicht kommen Sie in der Regel nicht in Berührung. Sie wird vom Treiber Ihrer Hardware bereitgestellt. Die Netzzugangsschicht verpackt die von den darüber liegenden Layern gelieferten Daten in "Hüllen" und versieht sie mit für das verwendete Netzwerk - etwa Ethernet - geeigneten physikalischen Adressierungen.

Über der Netzzugangsschicht liegt die **Vermittlungsschicht**. Sie stellt die Basisfunktionen bereit, um Datenpakete über das Netz zu verschicken. Sie ist somit das Herzstück eines TCP/IP-Netzes. Unabhängig davon für wen die Daten bestimmt sind und zu welcher Anwendung sie gehören, bietet die Vermittlungsschicht einen festen Standard, um auch unterschiedlichste Daten einheitlich transportieren zu können. In einem TCP/IP-Netzwerk wird die Vermittlungsschicht durch das Internet Protocol (IP) umgesetzt. IP ist ein so genanntes "verbindungsfreies" Protokoll. Es tauscht also keine Kontrollinformationen mit der Gegenstelle im Netzwerk aus (Handshaking). Sie fragt also nicht, ob der Kommunikationspartner überhaupt bereit ist, Daten zu empfangen. In diesem Punkt hängt IP von den darüber liegenden Protokollen der Transportschicht ab. IP erfüllt wirklich nur die Aufgabe, Daten von der Transportschicht in einheitliche Datenpakete zu packen und diese an die Netzzugangsschicht weiterzuleiten und umgekehrt.

Den Verbindungsaufbau und die Datenübermittlung zwischen den Hosts übernimmt die **Transportschicht**. Bei der Transportschicht können verschiedene Protokolle genutzt werden. Je nachdem, welche Anforderungen an die Zuverlässigkeit, Geschwindigkeit etc, gestellt werden, kann ein adäquates Protokoll in der Transportschicht zum Einsatz kommen. Die bekanntesten und wichtigsten Protokolle sind das Transmission Control Protocol (TCP) und das User Datagram Protocol (UDP). Die **Anwendungsschicht** ist der **Layer**, mit der Anwender direkt arbeiten. In ihr agieren die Netzwerkprogramme, die über die Anwendungsprotokolle auf Dienste wie E-Mail oder Web zugreifen. In dieser Schicht herrscht die größte Artenvielfalt an Protokollen. Für fast jeden Zweck wurde ein speziell abgestimmtes Protokoll geschaffen.

Hier eine kleine Übersicht:

Protokoll	Abkürzung	Beschreibung
Hypertext Transfer Protocol	<b>http</b>	Protokoll für das Surfen mit dem Internet-Browser
File Transfer Protocol	<b>FTP</b>	Hiermit lädt man Dateien aus dem Internet herunter bzw. hoch
Simple Mail Transfer Protocol	<b>SMTP</b>	Dieses Protokoll dient zum Versenden von Mails
Post Office Protocol	<b>POP</b>	Mit POP holt man Mails aus dem Postfach beim Provider ab
Internet Message Access Protocol	<b>IMAP</b>	wie POP jedoch leistungsfähiger und moderner
Domain Name Service	<b>DNS</b>	Einer URL (Hostname) wird eine IP-Adresse zugewiesen
Server Message Block	<b>SMB</b>	Mit diesem Protokoll werden unter Windows Drucker und Dateien freigegeben
Network File System	<b>NFS</b>	Freigabe von Dateisystemen unter Unix
Network Terminal Protocol	<b>TELNET</b>	Remote einloggen auf Unix-Rechnern
Dynamic Host Configuration Protocol	<b>DHCP</b>	Automatische IP-Zuweisung an einen Client

Ein Beispiel für ein direkt vom Anwender verwendetes Protokoll ist HTTP, durch das das Browsen im Web möglich wird. Indirekt - und damit unbemerkt vom Anwender - wird beispielsweise DNS verwendet, durch das Hosts anhand ihres Namens im Internet gefunden werden. Wann immer eine

Anwendung Daten per Netz übermitteln will, verpackt es diese zunächst in ihr eigenes Protokoll. Ein Web-Browser zum Beispiel codiert die Daten in HTTP-Befehle. In der Transportschicht werden diese dann in TCP-Pakete umgesetzt und korrekt adressiert. Anschließend macht IP in der Vermittlungsschicht die Daten "Internet-tauglich". Die Netzwerkhardware, wie ein Modem, schickt die Daten dann auf die Reise.

Auf dem Server, der die Daten empfängt, läuft der Vorgang umgekehrt ab. Die Daten wandern nun den TCP/IP-Stapel in umgekehrter Reihenfolge wieder nach oben. Die Web-Server-Software ganz zu oberst erhält die für sie verständlichen HTTP-Befehle und interpretiert diese. Die 'Antwort', zum Beispiel eine angeforderte HTML-Seite, geht dann wieder auf dem umgekehrten Weg auf die Reise.

#### **Adressierung**

Wie erwähnt erhält jedes Gerät eine eindeutige Adresse. Bei TCP/IP (genauer IPv4) erhält jeder Computer eine 32 Bit (= 4 Bytes) lange Zahl, die IP-Adresse. Die Hosts werden also - einfach gesprochen - durchnummeriert. Die Zahlenkolonne eines 32-Bit-Werts ist nicht gerade sehr übersichtlich und schon gar nicht leicht im Kopf zu behalten. Aus diesem Grund hat sich die "Punktnotation" eingebürgert. Statt einen 4 Byte umfassenden Wert in einem Zug durchzuschreiben, wird er in vier einzelne Werte gegliedert. Jeder dieser Werte stellt eines der vier Bytes dar. Jede dieser vier Zahlen liegt damit nur zwischen 0 und 255. Vier solche Werte sind sehr einfach zu merken.

Ein Beispiel für eine solche IP-Adresse ist 217.112.101.17

Ein TCP/IP-Netz besteht nicht einfach nur aus einem einzigen großen Netzwerk. Netze gliedern sich in Subnetze, die wie Hosts zweifelsfrei identifiziert werden müssen. Sie benötigen also auch eine Nummer. Eine IP-Adresse beinhaltet daher nicht nur eine Host-Nummer, sondern auch eine Zahl für das Netzwerk. Sie unterteilt sich in Bereiche für Host-Adresse und Netzwerkadresse. Die Aufspaltung ist dabei gar kein Hexenwerk. Es wird einfach eine Grenze festgelegt, wie viele Bits der 32 Bits langen IP-Adresse auf die Host- und wie viele auf die Subnetz-Adresse entfallen.

In einem TCP/IP-Netz darf jede IP-Adresse nur ein Mal vorkommen. Das Dilemma: Es gibt in lokalen Netzwerken viel mehr Computer, als es IP-Adressen für das Internet gibt. Früher oder später würden also die IP-Adressen knapp. Außerdem müssten Milliarden von Computern zentral verwaltet werden sowie diese direkt oder indirekt mit dem Internet verbunden sind.

Auch hier hilft das Subnetzkonzept. Der denkbare Adressraum wurde in zwei Bereiche aufgeteilt: In Adressen für das Internet und für das lokale Netzwerk. Die ersten Adressen werden zentral verwaltet und nur Computer, die direkt mit dem Internet verbunden sind, dürfen eine solche Adresse haben. Die Adressen, die fürs lokale Netzwerk reserviert sind, sind im Internet selbst ungültig. Damit können diese im LAN frei vergeben werden. Die Adressen, die für lokale Netzwerke reserviert sind, liegen in folgenden Bereichen:

10.0.0.0 bis 10.255.255.255

172.16.0.0 bis 172.16.255.255

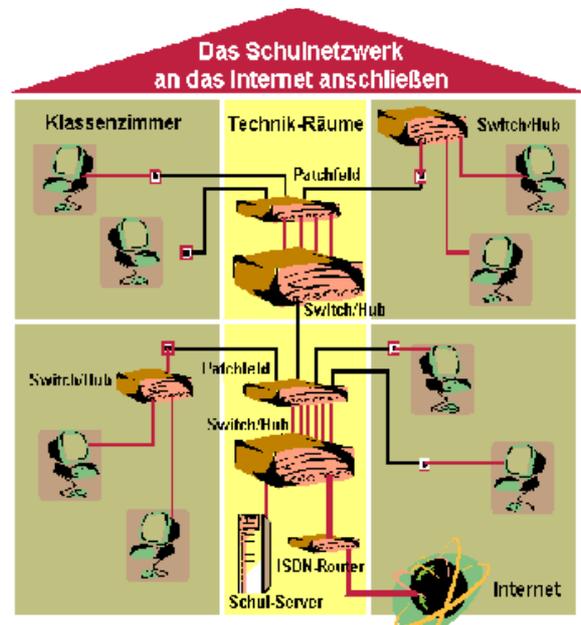
192.168.0.0 bis 192.168.255.255

Alle anderen IP-Adressen sind für das lokale Netzwerk tabu. Die Computer aus dem lokalen Netzwerk müssen dabei über Gateways und Router mit dem Internet kommunizieren, die den Datentransfer von einem Subnetz ins andere übersetzen.

Weitere Einschränkungen beim Vergeben von IP-Adressen an Hosts sind die reservierten Adressen. Sind alle Bits des Host-Anteils auf 0 gesetzt, beschreibt die entstehende Adresse keinen Host, sondern bezeichnet das gesamte Netzwerk. 127.16.0.0/16 verweist auf das Subnetz 127.16.\*. Die Host-Adresse 255 ist ebenfalls reserviert. Sie bezeichnet die Broadcast-Adresse, durch die alle Hosts im Subnetz auf einmal adressiert werden können. 192.168.0.255/16 bezeichnet beispielsweise die Broadcast-Adresse von Subnetz 192.168.\*. Die Adressbereiche 0.0.0.0/8 und 127.0.0.0/8 sind ebenfalls für bestimmte Aufgaben reserviert und daher nicht verfügbar.

## Ports

Bei TCP/IP können auf einem Host mehrere Services laufen. So kann ein Web-Server ebenso aktiv sein, wie ein Mail-Server oder ein Dienst für Netzwerkdrucker. Die einzelnen Dienste werden dabei über Ports~ bzw. Portnummern identifiziert. Diese Nummern sind 16-Bit-Werte, die angeben, welches Programm für die Verarbeitung der Anfrage zuständig ist. Eine TCP/IP-Kommunikation hat dabei immer einen Quellport und einen Zielport. Der Quellport identifiziert das Programm, das die Anfrage sendet. Dies kann etwa ein Port sein, den der Webbrowser (Port 80) nutzt. Der Zielport gibt hingegen den Port des Hosts an, der die Anfrage abarbeitet. Auf diesem Port läuft dann beispielsweise der Web-Server. Die Portnummern unter 256 sind für so genannte Well-Known-Services, wie FTP (Port 21) oder TELNET, reserviert. Die Werte zwischen 256 und 1024 werden für systemspezifische Dienste genutzt. Alle anderen über 1024 sind frei nutzbar. Sie dienen zum Beispiel für ausgehende Verbindungen als Quellports. Ports sind äußerst wichtig im Securitybereich und müssen beobachtet werden. Sie sind nicht zuletzt Objekt der Begierde von Angreifern. Das ist nicht weiter verwunderlich, wenn man bedenkt, dass die Ports für den Aufbau von Verbindungen zu verschiedenen Diensten notwendig sind. Diese Dienste sind aber die Eintrittstür für den Zugriff auf einen Rechner. Ein erster Hinweis auf einen akuten Angriff ist daher ein sogenannter Portscan. Ein Hacker lässt dabei eine Software (Port-Scanner) auf den Computer los, die durchtestet, welche Ports offen sind. Solche Angriffe sind nur durch korrekt konfigurierte Firewalls abzuwehren.



## Anbindung an das Internet

Das Internet ist ein weltweites Netzwerk, der Zugang wird über einen Provider verwirklicht, z.B. T-Online oder AOL. Um nun eine Verbindung vom Schulnetz zum Provider herzustellen, wird ein zusätzliches Gerät, ein sog. Router, benötigt. Dieser Router kann ein Hardware-Router sein, der nur diese eine Aufgabe wahrnimmt oder ein PC mit ISDN-Karte und spezieller Software (Software-Router), der zusätzlich noch weitere Aufgaben erledigen kann.

**Hub's** sind für die physikalische Verbindung zuständig. Ein **Switch** erledigt die gleichen Aufgaben wie ein Hub, verfügt aber unter anderem zusätzlich über die Fähigkeit ankommende Datenpakete nur an den Empfänger weiter zu leiten. Ein Hub dagegen sendet die Daten an alle angeschlossenen Rechner, wobei nur ein Rechner die Daten auch empfangen kann.



## Dienste im Internet

Über das Internet können Pakete für unterschiedliche Dienste verschickt werden. Entsprechende unterschiedliche Programme nutzen dies für den Anwender.

### ♦ WWW

Es ist in erster Linie das World Wide Web (WWW), das wesentlichen Anteil an der explosionsartigen Ausbreitung des Internet besitzt.

- An seinem Rechner arbeitet der Benutzer mit einem Browser (Web-Client). Dieser interpretiert Dateien, die im HTML-Format vorliegen und zeigt sie an.
- Der Browser baut über das Internet eine Verbindung mit dem Web-Server auf. Zur Kommunikation dient das http-Protokoll (Hypertext Transfer Protocol), das auf dem TCP/IP-Protokoll aufsetzt.

### ♦ Email

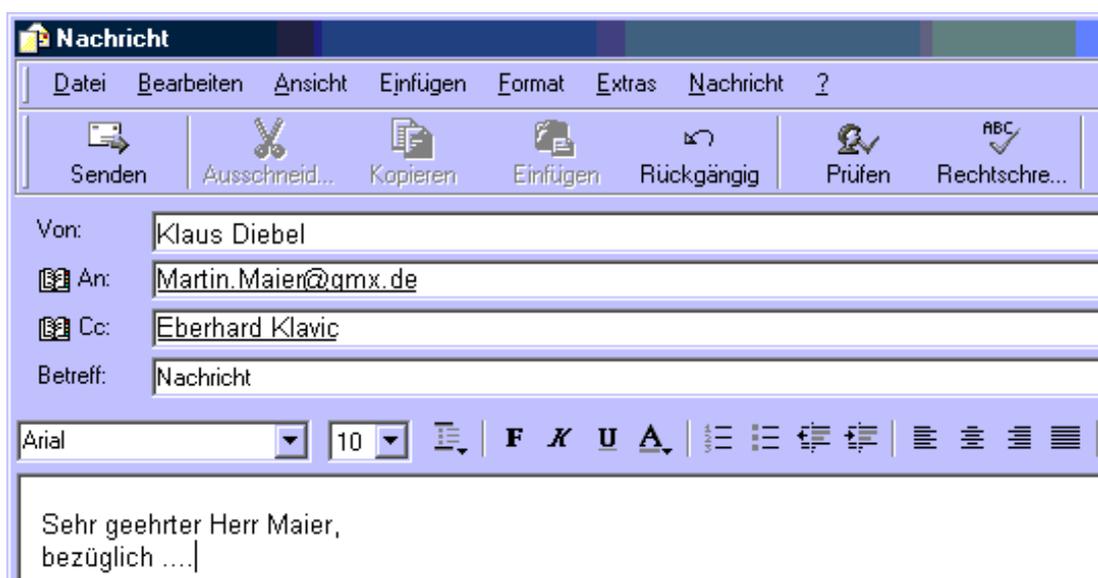
Ein Hauptnutzen des Internetzugangs ist die Möglichkeit, E-Mail (Electronic Mail) in die ganze Welt zu senden und zu empfangen. Die elektronische Post im Internet ist nichts anderes als das elektronische Gegenstück zur Papierpost.

Sie ermöglicht das Erstellen, das Adressieren und das Versenden von Nachrichten über die Netzleitungen des Internet. Man kann diese elektronischen Briefe öffnen, lesen und danach ablegen, später öffnen oder sogar ungelesen wegwerfen.

Ein augenscheinlicher Unterschied zur Papierpost liegt in der Geschwindigkeit der Zustellung. Verschickt man eine E-Mail, so ist die Entfernung zum Empfänger nahezu unerheblich, er kann sie kurz darauf entgegennehmen. Ebenso sind die anfallenden Kosten deutlich geringer als die Gebühren für die Papierpost. Auch gegenüber dem Telefon weist die E-Mail einen wesentlichen Vorteil auf: Der Empfänger der Post muss zum Übermittlungszeitpunkt nicht direkt anwesend sein (Zeitverschiebung, Arbeitszeit), er nimmt die E-Mail zu einem ihm geeigneten Zeitpunkt entgegen. Bei jeder E-Mail muss die Adresse des Empfängers angegeben werden. Typische Mailadressen im Internet haben die Form

*Nutzername@Mailserver (z. B. X.Huber@maildienst.de)*

Eine E-Mail wird mit einem speziellen E-Mail-Programm erstellt. Jedes E-Mail-Dokument besteht aus mehreren Teilen. Das E-Mail-Programm sieht für diese Teile feste Eingabefelder vor.



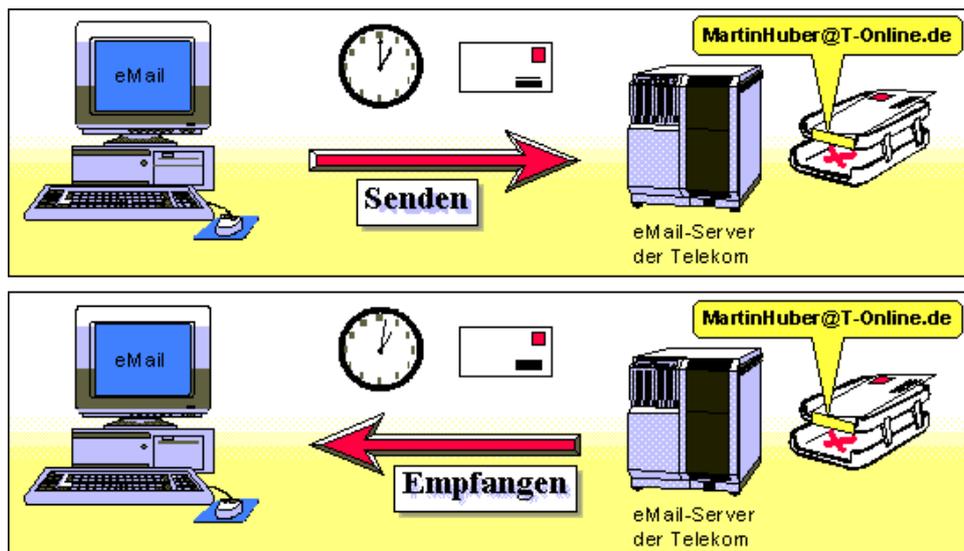
Wesentliche Teile einer E-Mail sind:

- Empfängeradresse („To:“): E-Mail-Adresse des Empfängers
- Kopie an („CC:“, carbon copy, Durchschlag): ggf. die E-Mail-Adresse einer Person (oder mehrerer), die eine Kopie der E-Mail erhalten soll; kann entfallen
- Absenderadresse („From:“): Wird meist automatisch vom E-Mail-Programm eingefügt
- Absendedatum und -zeit („Date:“): Wird automatisch angefügt
- Betreff Angabe („Subject:“): Jede E-Mail sollte eine kurze, treffende Angabe über den Inhalt bekommen. So kann der Empfänger sehr schnell den Grund der Nachricht erkennen, ohne diese vollständig lesen zu müssen. Da der E-Mail-Transport im Internet über viele Rechner mit unterschiedlichen Betriebssystemen läuft, sollte man im Feld Subject nur reine ASCII-Zeichen, d. h. keine Umlaute verwenden.
- Inhalt: Hier wird der eigentliche Textinhalt der E-Mail eingetragen.  
Mail-Anhang („Attachment“): Natürlich kann man an eine elektronische Post auch Anlagen oder ganze Dateien anheften. Über einen Dialog im E-Mail-Programm wählt man die gewünschte Datei in üblicher Weise aus.

Das Internet-Mail-System setzt das **Store- and Forward-Verfahren** ein. Der Sender überträgt die E-Mail zu seinem zugehörigen Mail-Server (dieser steht meist beim Provider; entspricht dem Mailserver in der E-Mail-Adresse des Senders). Bei diesem wird die Post zunächst gespeichert und in bestimmten Zeitintervallen, oder wenn der nächste Knotenrechner im Postweg sich vorher meldet, zum nächsten E-Mail-Knotenrechner übertragen. Dies kann über einige Stationen laufen, bis die Nachricht am Ziel-Mailserver ankommt. Die Kette der durchlaufenen Mail-Knoten ist teilweise auf dem vollständigen E-Mail-Kopf für den Empfänger ersichtlich. Nach der erfolgreichen Übertragung löscht jeder Zwischenknoten in der Postkette die temporär gespeicherte Nachricht.

Der Zielpostknoten (Ziel-Mailserver) speichert die Nachricht so lange, bis sie vom Empfänger abgeholt wird oder bis ein vorgegebenes Zeitlimit überschritten wird.

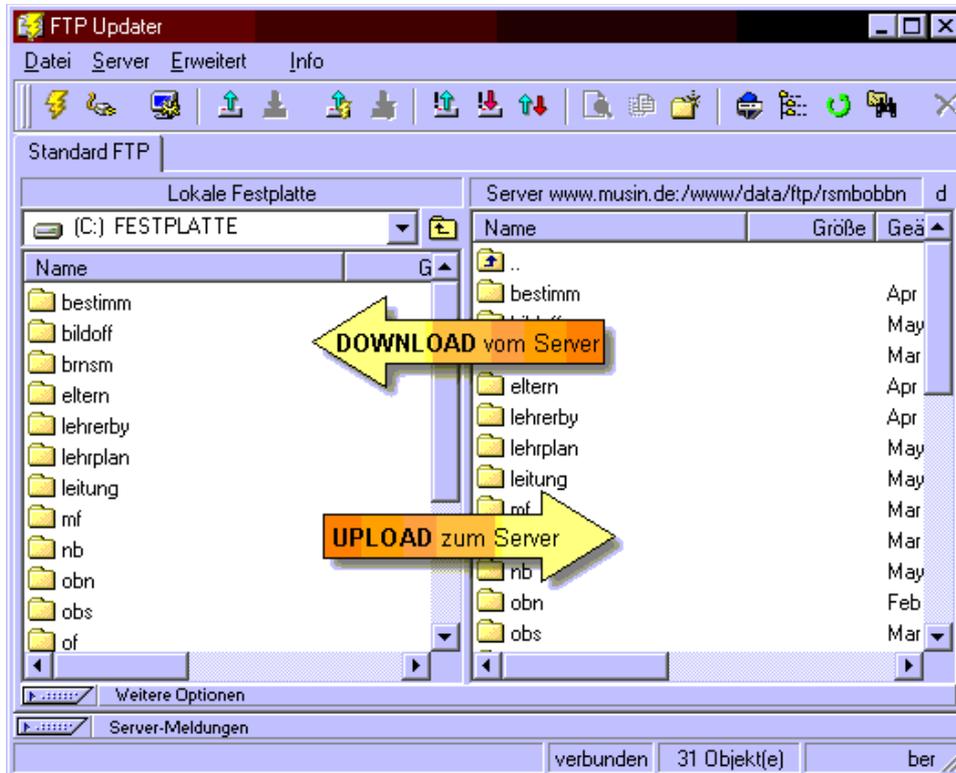
Die nachfolgende Abbildung zeigt das Senden und Empfangen von eMails:



#### ♦ FTP

Der FTP-Dienst (File Transfer Protocol) ermöglicht es, Dateien zwischen zwei Rechnern auszutauschen. Den FTP-Dienst benutzt man in zwei Hauptfunktionen:

- direkter Dateiaustausch zwischen zwei Systemen; hierzu muss man auf dem entfernten System eine Benutzerkennung und das entsprechende Passwort besitzen.
- anonymes FTP. Hier hat man Zugriff auf ein großes Angebot an Freeware, Shareware und Updatedateien, die man sich kostenlos von den FTP-Servern downloaden darf. Zur Anmeldung dient als Benutzername *anonymous* und als Passwort wird die E-Mail-Adresse des Anwenders eingesetzt. Meist ist die Anzahl der *Anonymous-User*, die gleichzeitig bei einem FTP-Server Zugang haben, auf eine feste Zahl begrenzt.



Mit dem FTP-Dienst lassen sich Dateien zwischen den Rechnern in beiden Richtungen austauschen, sofern der Benutzer hierzu die Berechtigung hat.

#### ◆ Chat

Es handelt sich um Unterhaltung per Tastatur weltweit zum lokalen Internet-Zugangstarif (Vorstufe des Telefonierens im Internet). Notwendig ist ein Chat-Server, auf dem sich zwei oder mehrere (Konferenz-) Teilnehmer in einem "channel" zusammenfinden.

#### ◆ News

Nach Art eines schwarzen Bretts können Texte und Bilder abgelegt werden, die dann von anderen gelesen werden können (sollen).

### **Informationen suchen und finden im Internet**

Um mit einem geringen Zeitaufwand an gewünschte Informationen zu gelangen, sind gewisse Grundkenntnisse über den Aufbau, die Ressourcen bzw. Informationsquellen, die Informationsstruktur und die vielfältigen Möglichkeiten des Einsatzes von Hilfsmitteln zur gezielten Suche im Internet notwendig.

„Wer sucht, der findet!“ Dies lässt sich leicht aussprechen, aber bei der Umsetzung dieses Sprichwortes treten Schwierigkeiten auf, besonders dann, wenn man ohne die richtige Systematik an die Aufgabenstellung der Suche herangeht.

Zu Beginn einer Recherche im Internet sollte ein Anforderungsprofil für die Suche nach den gewünschten Informationen erstellt werden.

- Wer sucht die Information?
- Welche Information wird gesucht?

- Wo könnten die Informationen vorhanden sein?
- Wie detailliert sollen die Informationen beschrieben werden?
- Wie schnell sollen erste Ergebnisse vorliegen?
- Welchen Umfang soll die Suche haben?

### Prinzipielle Arbeitsweise von Suchmaschinen

Die Such- und Sammelmaschinen werden als „Robots“ bezeichnet. Sie nutzen die Verweisstrukturen des WWW, die von einem Dokument ausgehen. Die Suchrobots gehen dabei von Dokumenten aus, die ihnen bekannt sind, und analysieren die darin enthaltenen Verweise zu weiteren Dokumenten. Von den Suchrobots wird eine lexikalische Analyse durchgeführt, bei der die Inhalte der Dokumente für die Aufnahme in die Datenbank aufbereitet werden. Zwischen den von den Suchmaschinen eingesetzten Techniken für die Handhabung der riesigen Datenmengen bestehen bezüglich Analyse und Aufbereitung große Unterschiede.

Die Aufgaben einer Suchmaschine werden prinzipiell in mehrere Teilbereiche aufgliedert:

- Einsammeln der Web-Dokumente
- Analyse und Indexierung der Web-Dokumente
- Regelmäßige Kontrolle der bereits erfassten Web-Dokumente und Aktualisierung der Datenbank
- Suchanfragen bearbeiten
- Einsammeln der Web-Dokumente

Suchmaschinen verwenden häufig Kataloge oder andere Suchmaschinen, um den URL zu ermitteln, der für den Start der automatischen Suche erforderlich ist. Das Einsammeln der Web-Dokumente erfolgt aus den Informationsquellen des WWW, den News-Gruppen und den FTP-Servern.

Grundsätzlich werden folgende Informationsquellen im Web nicht erreicht:

- Web-Dokumente, die bei den Suchmaschinen nicht eingetragen sind bzw. auf die kein Verweis existiert.
- Auf den Web-Servern existierende Dateien, die Informationen enthalten, die von den Suchmaschinen nicht erfasst werden sollen
- Geschützte und dynamisch generierte Web-Dokumente, die durch Passwortregistrierung, Firewall bzw. über Formulare dynamisch generiert werden
- Dateien, die außerhalb der anonymus-FTP-Server erreichbar sind.

### Analyse und Indexierung der Web-Dokumente

Das Angebot von Suchmaschinen und Suchoperationen ist stark von der Indizierung der eingesammelten Web-Dokumente und der daraus generierten Datenbank abhängig. Die meisten Suchmaschinen indizieren die Web-Dokumente im Volltext nach sogenannten Stockwortlisten. Für den Aufbau eines Teilindex indizieren Suchmaschinen häufig nach dem Titel-Tag oder nach den Heading-Tags. Die HTML (Hypertext Markup Language) bietet mit dem Meta-Tag dem Web-Autor die Möglichkeit, Deskriptoren und Zusatzinformationen zum Web-Dokument einzutragen. Die Suchmaschinen, die Meta-Tags unterstützen, indizieren die dort enthaltenen Informationen.

### Aktualisierung

Die von Suchmaschinen erfassten URLs werden in zeitlichen Abständen aktualisiert. Das Übertragungsprotokoll im WWW (http) enthält mit dem http-Request einen Mechanismus, bei dem die Übertragung der letzten Änderung erfolgt. Das Web-Dokument wird nur dann zur Indizierung und Aktualisierung der Einträge in der Suchmaschinen-Datenbank übertragen, wenn über den http-

Request festgestellt werden kann, dass seit der letzten Indizierung eine Änderung am Web-Dokument erfolgt ist.

### Suchanfragen bearbeiten

Die von Suchmaschinen angebotenen Methoden oder Operationen hängen vom Verfahren der lexikalischen Analyse und Indizierung der Web-Dokumente ab.

### Vorbereitung der Suche

Jede Suche unterscheidet sich von anderen je nachdem, ob man nur eine spezifische Information zu einem Begriff wünscht oder ob eine komplexe Recherche zu diesem Begriff durchgeführt werden soll. Weiterhin ist von Bedeutung, inwieweit sich die Suchabfrage präzisieren lässt oder ob für die Suche eine Erweiterung des Suchraums, also die Einbeziehung eines übergeordneten Themengebiets erforderlich ist. Dadurch erhält man einen möglicherweise stark erweiterten Datenbestand, dessen Verifizierung mit einem großen Aufwand verbunden ist. Die Erstellung des jeweiligen Anforderungsprofils hat eine nicht zu unterschätzende Bedeutung. Ein weiterer Aspekt des Anforderungsprofils ist die Erleichterung der Auswahl eines geeigneten Suchwerkzeugs.

### Beispiel 1:

Aufgabe: Die Beschaffung von Informationen zum Thema „NASA“.

### Fragestellungen:

- Ist nur die Begriffserklärung von NASA gewünscht?
- Oder möchte man Informationen zu Projekten, wissenschaftliche Thesenpapiere?

### Beispiel 2:

Aufgabe: Eine Software wird gesucht.

- Welche Problemstellung soll die Software lösen?
- Betriebssystem, spezielle Software?

### Das Anforderungsprofil der Suche

Ein guter Ansatz, das Anforderungsprofil der jeweiligen Suche zu erarbeiten, ist die Beantwortung der im untenstehenden Merksatz implizierten Fragestellungen:

Wer sucht Was für Wen, Wo, Wie, Womit, Wann, in Welchem Umfang und Warum?

#### a) Was wird gesucht?

- Eine spezifische Einzelinformation
- Alle verfügbaren Informationen zum Suchbegriff
- Eine Begriffserklärung
- Eine wissenschaftliche Veröffentlichung (Diplomarbeit, Dissertation, Forschungsbericht)
- Eine Firma
- Ein Softwareprodukt - Für welche Hardware? / Für welches Betriebssystem?
- Sonstige Fragestellungen

#### b) Wer sucht?

- Mann, Frau, Kind
- Welche allgemeine Qualifikation hat der Suchende?
- Schulische Ausbildung
- Welchen Beruf hat der Suchende?
- Sonstige Fragestellungen

c) Für Wen soll gesucht werden?

- Wie oben

d) Wo wird gesucht?

- Ort des Suchbeginns (An einem Rechenzentrum - welches? Bei einem Provider - welcher?)
- Welche Ressourcen sollen bei der Suche berücksichtigt werden? (Bibliotheken - welche? Spezielle Datenbanken? Anwaltssuche? Patente?)
- Geographische Zuordnung für die Suche (weltweit, USA, ...)

e) Wie soll gesucht werden?

- Organisationsmethodik der Suche  
Bereitstellung von Ressourcen (wie viele Personen werden bei der Suche eingesetzt?)  
Wahl des Ausgangspunktes für die Suche (Person A sucht im europäischen Raum, Person B sucht in der USA ...)
- Einbringen von Methoden der Suchverfeinerung  
Volltextsuche, Schlüsselwortsuche, Wildcardsuche ....

f) Warum erfolgt die Suche?

- Zur Erweiterung des eigenen Wissenstandes zum Thema
- Für die Erstellung eines Vortrages
- Die gefundenen Informationen sind für ein Buchprojekt
- Sonstige Zwecke

Suchen mit einem Katalog

Kataloge sind im Grunde nichts anderes als riesige Link-Listen und Bookmark-Verzeichnisse.

Bei diesem Angebot sitzen Menschen da und tun den ganzen Tag nichts anderes als sich Internet-Seiten anzusehen, sie zu bewerten und – die für gut befundenen Seiten in den Katalog aufzunehmen.

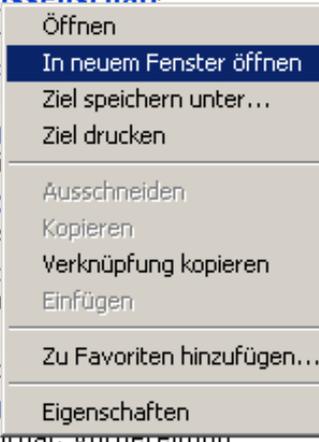
Rufen Sie den Link mit der rechten Maustaste auf und wählen 'Sie im neuen Fenster öffnen', so bleibt Ihnen die Seite mit den Kategorien und Links erhalten.

The screenshot shows the WEB.DE website interface. At the top, there is a navigation bar with links for 'Suche', 'Themen', 'Dienste', 'FreeMail', 'Hilfe', and 'Kontakt'. Below this, there is a promotional banner for 'WEB.DE FreeMail' with the text 'Nur hier: WEB.DE FreeMail Testsieger bei Stiftung Warentest Jetzt registrieren!'. The main content area features a search bar with a 'Suchen' button and a 'Live-Suche' link. Below the search bar, there are radio buttons for 'Suche in: Web-Seiten', 'Nachrichten', 'Aktienkursen', and 'Newsgroups'. A 'Dienste' (Services) section is displayed as a grid of links, including categories like 'Kommunikation', 'Aktuelles', 'Unterhaltung', 'Service', and 'Verbraucher'. A 'Neu' (New) section at the bottom highlights 'Fußballtipp' and 'Specials: Popkomm, Funkausstellung 2001'.

Dienste					<a href="#">mehr ...</a>
Kommunikation	Aktuelles	Unterhaltung	Service	Verbraucher	
FreeMail	Nachrichten	MillionenKlick	Routenplaner	<b>Shopping</b>	
SmartSurfer	<b>Börse</b>	Grußkarten	Staumelder	paybox	
FreeSMS	Sport	Comics	Baustellen	Stromtarife	
<b>WAP</b>	<b>Formel1</b>	Surftipps	<b>Download</b>	Telefontarife	
Newsgroups	Wetter	MP3	Übersetzer	Gratis	
<b>Kalender</b>	Gesundheit	Radio	Webführersch.	Kataloge	
Chat	Horoskop	<b>Handy-Logos</b>	Externe	Branchenbuch	
TrustCenter	TV-Programm	S.of Avalon	DB Fahrplan	<b>Dienstleister</b>	
Weckruf			Flugtickets		

**Neu:** Fußballtipp Specials: Popkomm, Funkausstellung 2001

Themen		Themen	
<b>Auskunft &amp; Information</b> Telefon, Internet, Strom	<b>Auto &amp; Verkehr</b> Gebrauchtwagen, Mo	<b>Auskunft &amp; Information</b> Telefon, Internet, Strom	<b>Auto &amp; Verkehr</b> Gebrauchtwagen, Mo
<b>Bauen &amp; Wohnen</b> Immobilien, Garten, Renovieren	<b>Beruf &amp; Karriere</b> Aus- & Weiterbildung	<b>Bauen &amp; Wohnen</b> Immobilien, Garten, Renovieren	<b>Beruf &amp; Karriere</b> Aus- & Weiterbildung
<b>Bildung &amp; Wissenschaft</b> Schule, Studium, Wissenschaft	<b>Computer &amp; Software</b> Hardware, Software	<b>Bildung &amp; Wissenschaft</b> Schule, Studium, Wissenschaft	<b>Computer &amp; Software</b> Hardware, Software
<b>Einkaufen &amp; Sparen</b> Online einkaufen, Dienstleistungen	<b>Freizeit &amp; Unterhaltung</b> Für Männer, Kino, Games	<b>Einkaufen &amp; Sparen</b> Online einkaufen, Dienstleistungen	<b>Freizeit &amp; Unterhaltung</b> Für Männer, Kino, Games
<b>Gesellschaft &amp; Politik</b> Familie, Religion, Parteien	<b>Gesundheit &amp; Medizin</b> Beauty, Wellness, Nahrung	<b>Gesellschaft &amp; Politik</b> Familie, Religion, Parteien	<b>Gesundheit &amp; Medizin</b> Beauty, Wellness, Nahrung
<b>Internet &amp; Kommunikation</b> Breitband, Festnetz, Mobilfunk	<b>Kunst &amp; Kultur</b> Literatur, Musical, Netze	<b>Internet &amp; Kommunikation</b> Breitband, Festnetz, Mobilfunk	<b>Kunst &amp; Kultur</b> Literatur, Musical, Netze
<b>Musik &amp; Medien</b> Multimedia, Presse, Radio & TV	<b>Private Homepage</b> Tipps & Tricks für Webmaster	<b>Musik &amp; Medien</b> Multimedia, Presse, Radio & TV	<b>Private Homepage</b> Tipps & Tricks für Webmaster
<b>Sport &amp; Fitness</b> Formel1, Fußball, Tennis	<b>Städte &amp; Regionen</b> Berlin, Hamburg, Karlsruhe	<b>Sport &amp; Fitness</b> Formel1, Fußball, Tennis	<b>Städte &amp; Regionen</b> Berlin, Hamburg, Karlsruhe
<b>Touristik &amp; Reisen</b> Angebote, Journal, Vorbereitung	<b>Wirtschaft &amp; Finanzen</b> Geld, Steuern, Versicherungen	<b>Touristik &amp; Reisen</b> Angebote, Journal, Vorbereitung	<b>Wirtschaft &amp; Finanzen</b> Geld, Steuern, Versicherungen



### Suche mit einer lokalen Suchmaschine

Viele Angebote im Internet ermöglichen es dem Anwender, innerhalb dieses Angebots gezielt nach Informationen zu suchen. Auch 'Web.de' bietet eine solche lokale Suchmaschine.

### Suche mit einer Suchmaschine

## Suche mit einer anderen Suchmaschine: www.google.de

Erweiterte Suche Einstellungen Preferences (English) Suchtipps

Google™ nasa Google-Suche Auf gut Glück!

Suche:  Das Web  Seiten in Deutsch  Seiten aus Deutschland

Seiten in **Deutsch** wurden nach **nasa** durchsucht. Resultate **1 - 10** von ungefähr **92,100**. Suchdauer: **0.09** Sekunden.

Kategorien: [World](#) > [Deutsch](#) > [Regional](#) > [Deutschland](#) > [Sachsen-Anhalt](#)

**NASA GmbH**  
 Fahrplanauskunft für den Personennahverkehr in Sachsen-Anhalt  
 Beschreibung: Fahrplanauskunft für den Personennahverkehr in Sachsen-Anhalt  
 Kategorie: [World](#) > [Deutsch](#) > [Regional](#) > [Deutschland](#) > [Sachsen-Anhalt](#)  
[www.nasa.de/](#) - 2k - [Im Archiv](#) - [Ähnliche Seiten](#)

**Nachrichten aus dem Südlichen Afrika**  
 Unsere Dokumentationsstelle liefert kompetent & kontinuierlich Hintergrundinformationen zum politischen und sozioökologischen Geschehen im Südlichen ...  
[www.nasa-basel.ch/](#) - 2k - [Im Archiv](#) - [Ähnliche Seiten](#)

**NASA RP 1398: Local Circumstances for Europe - Germany, ...**  
 Table 17. Local Circumstances for Europe - Germany Total  
 Solar Eclipse of 1999 August 11. ...  
[umbra.nascom.nasa.gov/eclipse/990811/tables/table\\_17.html](#) - 26k - [Im Archiv](#) - [Ähnliche Seiten](#)

**Das Portal zur NASA: News, Links, Bilder, ...**  
 Deutschsprachige Website mit Nachrichten über die Aktivitäten der NASA (auf englisch). Links zur NASA. Bilder mit Bezug zur NASA, hauptsächlich von ...  
[www.nasa-news.de/](#) - 2k - [Im Archiv](#) - [Ähnliche Seiten](#)

**NASA - Network**  
[www.nasa-productions.de/](#) - 2k - [Im Archiv](#) - [Ähnliche Seiten](#)

## Boolesche Abfragen

Im Feld Suche definieren Sie die Suche wie auch im normalen Suchmodus. Allerdings verwenden Sie hier Operatoren:

Ziel	Operator	Bedeutung	Beispiel	Ergebnis
alle Begriffe	AND &	sowohl ... als auch ...	George & Boole	alle Seiten, die beide Begriffe enthalten, unabhängig von ihrer Position im Text
irgendeinen der Begriffe	OR 	entweder ... oder ...	George   Boole	alle Seiten, die mindestens einen der Begriffe enthalten
diesen Begriff nicht	AND NOT !	alle ... außer ...	Boole ! George	alle Seiten, die 'Boole' und nicht 'George' enthalten
Mischform			Richard AND (König OR Löwenherz)	alle Seiten, die 'Richard' und entweder 'König' oder 'Löwenherz' enthalten

## Ergebnisse einer Suchmaschine

Die Ergebnisse werden von vielen Suchmaschinen durchnummeriert. Neben dem Seitentitel wird eine kurzer Textauszug der Seite angezeigt sowie die URL, das Datum der letzten Änderung, die Größe des Dokuments und die Sprache, in der es verfasst ist.

## Meta-Suchmaschinen

Der Begriff Meta kommt aus dem griechischen und bedeutet soviel wie „über“. Man sagt der Meta-Suchmaschine, in welchen Suchmaschinen sie nach einem Begriff suchen soll, und das Ergebnis ist die Quintessenz der einzelnen Trefferlisten. Meta-Suchmaschinen erfassen einen wesentlich größeren Teil des WWW als eine einzelne Suchmaschine.

## MetaGer, die MetaSuchmaschine über deutschsprachige Suchmaschinen



Ein Service des [RRZN&RVS](#), [Universität Hannover](#)

Geben Sie einfach ein oder mehrere Suchwörter ein:

mit internationaler Suche,  [Metacrawler](#):  [Dmoz](#):

Oder suchen Sie zunächst die richtigen Suchwörter?  
Dann gehen Sie dazu auf unseren [Assoziator](#).



Wenn Suchmaschinen nicht reichen:  
Fragen Sie die *Menschen* des [Meta-re-SearchTeams](#)

### Hilfe

[URL  
so/abmelden](#)

[Meta-ICQ-Test](#)

[Suchmaschinen  
Labor](#)

[Qualitätstest](#)

[Zugriffszahlen](#)

[Browserkrieg](#)

### Die von uns ausgewählten besten Suchdienste:

(Falls Sie Voreinstellungen unter diesen ausgewählten Diensten ändern wollen, klicken Sie bitte den entsprechenden Schalter an.)

- |   |   |   |  |  |
|---|---|---|--|--|
| <input checked="" type="checkbox"/> <a href="#">Netfind</a> | <input checked="" type="checkbox"/> <a href="#">Altavista</a> | <input checked="" type="checkbox"/> <a href="#">Yahoo</a> | <input checked="" type="checkbox"/> <a href="#">web.de</a>   | <input checked="" type="checkbox"/> <a href="#">AltaVista.de</a> |
| <input checked="" type="checkbox"/> <a href="#">Lycos</a>   | <input checked="" type="checkbox"/> <a href="#">QualiGO</a>   | <input checked="" type="checkbox"/> <a href="#">Witch</a> | <input checked="" type="checkbox"/> <a href="#">T-Online</a> | <input checked="" type="checkbox"/> <a href="#">Infoseek</a>     |
| <input type="checkbox"/> <a href="#">MSN</a>                | <input type="checkbox"/> <a href="#">Speedfind</a>            | <input type="checkbox"/> <a href="#">Walhello</a>         | <input type="checkbox"/> <a href="#">AllesKlar</a>           | <input type="checkbox"/> <a href="#">Intersearch.at</a>          |
| <input type="checkbox"/> <a href="#">Dmoz</a>               | <input type="checkbox"/> <a href="#">AllTheWeb</a>            | <input type="checkbox"/> <a href="#">Alexana</a>          | <input type="checkbox"/> <a href="#">Crawler.de</a>          | <input type="checkbox"/> <a href="#">Uni-Hannover</a>            |
| <input checked="" type="checkbox"/> <a href="#">nhf</a>     |   |   |  | <input type="checkbox"/> <a href="#">Online-Favoriten</a>        |

[ALLE Suchdienste absuchen](#)

Für detaillierte Anfragen empfehlen wir Ihnen die direkte Benutzung der obigen Suchdienste.

### Weitere Meta-Suchmaschinen

- [www.highway61.com](http://www.highway61.com)
- [www.ertools.ch](http://www.ertools.ch)
- [www.infosuche.com](http://www.infosuche.com)
- [www.infosuche.ch](http://www.infosuche.ch)
- [www.mp3bot.de](http://www.mp3bot.de)
- [www.metasuchdienst.de](http://www.metasuchdienst.de)
- [www.metafind.com](http://www.metafind.com)

### Sicherheit in Netzwerken

#### Gefahren

Mitte der Neunziger Jahre wurden mit der *Hyper Text Markup Language (HTML)* und dem dazugehörigen Übertragungsprotokoll *Hyper Text Transfer Protocol (HTTP)* die Grundlagen geschaffen, die aus dem Internet das World Wide Web machten, wie wir es heute kennen. Bald war der Wunsch da, die Nutzung des WWW komfortabler und interessanter zu machen und den Funktionsumfang zu erhöhen. Hier sind Java, JavaScript, ActiveX und auch Cookies als entscheidende Komponenten aufzuführen. Alle diese Technologien haben tatsächlich dazu beigetragen, das Netz lebendiger und besonders hinsichtlich der Interaktivität funktionaler zu gestalten. Sie haben aber in gleicher Weise Sicherheitsprobleme mit sich gebracht. So wird man sich in vielen Fällen zwischen Funktionalität und Sicherheit entscheiden müssen.

## Java

Von der Firma *Sun Microsystems* wurde Java 1995 erstmals vorgestellt und in der Computerszene sofort mit großer Begeisterung aufgenommen. Diese plattformunabhängige, objektorientierte Programmiersprache mit der C/C++ - ähnlichen Syntax ist vielseitig und mächtig: Grafik, Animationen, Audio, Video, grafische Benutzeroberflächen, Netzwerkzugriffe und vieles andere lassen sich in Java relativ schnell und einfach realisieren. Entscheidend für den Siegeszug dieser Programmiersprache war aber erstmals die Möglichkeit, über ein Rechnernetz Programme von einem fremden Rechner direkt in den Arbeitsspeicher des eigenen Rechners zu laden und dort automatisch auszuführen. Vorrangig sind hier die Java-Applets zu nennen, die stets in eine HTML-Seite eingebettet sind und daher von einem Java-fähigen Browser ausgeführt werden können. Dazu wird im Arbeitsspeicher ein bestimmter Bereich, die sogenannte Sandbox, zur Verfügung gestellt. Von dort aus haben die Applets nur sehr begrenzte Zugriffsrechte auf die Ressourcen des Computers und sind daher keine große Bedrohung. Jedoch können sie auch hier die Prozessorleistung stark beeinträchtigen und einen Rechner z.B. zum Absturz bringen. Um die Anwendungsmöglichkeiten zu erweitern, wird ein Java-Applet nicht immer in der Sandbox ausgeführt. Zur Verbesserung der Sicherheit wurde dann die Möglichkeit eingeführt, dass Applets über ein digitales Zertifikat verfügen. Solche als "trusted" eingestuft Applets können dann z.B. bestimmte Dateien auf der Festplatte lesen und neue erzeugen, oder vom Benutzer in ein Formular eingegebene Daten an den Server, von dem sie heruntergeladen wurden, zurücksenden. Auch müssen manche Applets z.B. temporäre Dateien schreiben, deshalb erhalten sie von den Benutzern Zugriff auf das Verzeichnis */tmp* unter der Annahme, dass sich ohnehin dort nichts Wichtiges befindet. Leider führen die meisten Editoren temporäre Verzeichnisse von Dokumenten und e-Mail genau dort, so dass bössartige Applets diese Dateien kopieren können und versuchen können, sie übers Netz in böser Absicht weiterzugeben. Als Abhilfe den Zugang diesbezüglich zum Netz zu unterbinden, würde sogleich die Funktionalität der Applets stark einschränken. Betrachtet man alles zusammen, so kann man sagen, dass es sich bei Java um ein eher sicheres System handelt. Allerdings sollte man sich von Fall zu Fall entscheiden, ob man Applets einsetzen möchte oder nicht.

## JavaScript

JavaScript wurde von Netscape entwickelt und hat mit Java nur die Namensähnlichkeit gemeinsam. Es handelt sich bei JavaScript um eine Skriptsprache, die direkt in den HTML-Code hineingeschrieben wird. Aufgrund des entsprechenden Tags (`<script language = "JavaScript"> </script >`) erkennt der Webbrowser den Beginn bzw. das Ende des Skripts. Das JavaScript wird also direkt beim Seitenaufbau von der Interpreter-Software des Webbrowsers bearbeitet und muss nicht extra herunter geladen werden wie ein Applet. Da JavaScript die Funktionalität von Webseiten wesentlich erhöht, würde ein Verzicht darauf sogleich deutlich spürbar. In der Vergangenheit sind immer wieder in beiden gängigen Browsern Sicherheitslücken bekannt geworden. Die Auswirkungen bestanden im Ausspähen der persönlichen Daten wie e-Mail-Adressen oder gespeicherten Passwörtern, im Versenden von Nachrichten und im Auslesen von Dateien auf der Festplatte. Neuere Versionen der Browser haben diese Lücken erkannt und behoben. Dennoch bestehen z.B. Probleme wie

- Manipulierbarkeit des Aussehens des Browsers.
- Erzeugen von Fenstern, die scheinbar zu anderen Anwendungen gehören, und dadurch falsche Informationen anzeigen und Eingaben in Formulare abfangen.
- Die Möglichkeit des Durchschleusens (»tunneln«) von anderen aktiven Inhalten (z.B. ActiveX).

Der Umgang mit Benutzernamen, Kennwörtern und Kreditkartennummern sollte daher stets vorsichtig geschehen. Interessante Angriffsziele sind: Das Internet-Banking und die Online-Börse.

## ActiveX

Die Antwort von Microsoft auf Java und JavaScript war ActiveX. Mit ActiveX können Webseiten interaktiv gestaltet werden, allerdings nur, wenn der Microsoft Internet Explorer verwendet wird. Ähnlich wie Java-Applets arbeitet auch ActiveX mit Elementen, hier Controls genannt, die entweder auf dem System des Nutzers bereits vorhanden sind und dann von der darzustellenden Webseite einfach nur aufgerufen werden, oder zusätzlich zur dargestellten Webseite heruntergeladen und anschließend automatisch installiert werden. Dabei können ActiveX-Controls frei auf alle Systemressourcen zugreifen. Sie können u.a. Dateien erzeugen, öffnen, ändern oder übers Internet verschicken. Auch die *registry*, die ja bekanntlich Informationen über Systemeinstellungen, angemeldete Benutzer und vieles mehr enthält, ist vor Veränderungen durch ActiveX-Controls nicht geschützt.

## Viren, Makroviren, E-mail-Viren

### Malicious Software („böswillige Software“) oder kurz Malware

Ein Virus ist definiert als ein Programm, das sich selbständig vermehren und verbreiten kann. Schreibt sich das Virus z.B. in das Autostart-Verzeichnis, wird es bei jedem Start des Rechners mit aktiviert. Manche Viren schreiben sich auch in den Master-Boot-Sektor, auf den der Rechner bei jedem Einschalten zurückgreifen muss. Diese Viren sind besonders gefährlich, da auch ein Virens Scanner eine minimale Systemumgebung benötigt und somit höchstens gleichzeitig mit dem Virus gestartet werden kann. Nach erfolgreicher Vermehrung führen die Viren weitere, oft lästige, aber harmlose Aktionen durch wie Änderungen der Farbeinstellungen, oder Ausgeben von Text am Bildschirm. Systemabstürze oder das automatische Herunterfahren des Rechners sind oft schon folgenreicher, da ein Datenverlust dabei nicht immer ausgeschlossen werden kann. Gefährlich wird ein Virus dann, wenn es versucht, die persönlichen Daten auszuspähen wie z.B. die Anmeldeinformationen für den Internetzugang, Kreditkartennummern oder Passwörter. Aus diesem Grund sollte man auch Passwörter nie auf der Festplatte abspeichern, da die Verschlüsselung von Windows mit im Internet frei erhältlicher Software entschlüsselt werden kann. Darüber hinaus ist ein Wechsel der Kennwörter in unregelmäßigen Abständen anzuraten. Die Eingabe persönlicher Daten im Internet, insbesondere von Kreditkartennummern oder Ähnlichem, sollte deshalb nur über sichere *HTTPS*-Verbindungen (*Secure Hyper Text Transfer Protocol*) erfolgen. Die Adressen dieser Verbindungen sind an dem Beginn mit *https://* zu erkennen. Die größte Bedrohung besteht jedoch durch Viren, die in der Lage sind, ganze Festplatteninhalte zu löschen oder sogar das BIOS zu überschreiben.



Makroviren sind in Visual Basic geschrieben und befallen Dokumente der Microsoft Office Familie. Da Visual Basic auf nahezu jede Systemressource Zugriff hat, ist es möglich, dass diese Viren ähnlich gefährlich werden wie herkömmliche Viren. Deshalb sollte man nur Dokumente aus vertrauenswürdiger und virenfreier Umgebung mit aktivierten Makros öffnen. Bei Deaktivierung kann immer unter dem Menüpunkt *Extras/Makro/Makros...* die Liste der Makrodateien angesehen werden, was allerdings einen Virenbefall nicht generell ausschließt. In den Sicherheitsstufen der Office-Anwendungen sollte daher im Normalfall die Stufe *Mittel* gewählt werden, um auf das Vorhandensein von Makros aufmerksam gemacht zu werden.

E-Mail-Viren sind eine Weiterentwicklung von Makroviren. Vor allem sind hier die in VB-Script geschriebenen Viren zu nennen, die sogar bereits bei der Ansicht der Vorschau aktiv werden. Diese Viren, die vom Netscape Navigator grundsätzlich nicht ausgeführt werden, werden auch als Würmer bezeichnet, da sie – im Gegensatz zu "echten" Viren - keine Dateien infizieren oder sich auf der Festplatte einnisten. Ein Wurm ist eigentlich nur ein Schadprogramm, das sich von Computer zu Computer via Netzwerk selbsttätig weiter verbreitet. Die 'Absicht' der Würmer ist es, so viele Computer wie möglich innerhalb eines Netzwerks zu befallen. Würmer brauchen, sind sie erst einmal auf den Weg gebracht, kein menschliches Zutun, um sich rasend schnell innerhalb eines

Firmennetzwerks oder über das Internet zu verbreiten. Sie benutzen beispielsweise die E-Mail-Funktionen eines Rechners, um sich an beliebige Internetadressen zu versenden. Und sie sind eher einfach zu programmieren. Für den Loveletter dürften das Microsoft-Handbuch (VBScript), ein Nachmittag und eine ordentliche Portion kriminelle Energie genügt haben, um einen Schaden von geschätzten 2,5 Milliarden Dollar weltweit anzurichten! Daher sollte man immer im Umgang mit Datei-Anhängen höchste Vorsicht walten lassen.

## Trojaner

Um über das Internet einen fremden Computer auszuspionieren, muss man nur sogenannte Backdoor-Programme oder Trojaner einsetzen. Dies ist Software, die speziell für diesen Einsatz konzipiert wurde und relativ einfach zu bedienen ist. Daher stellen Trojaner eine ernstzunehmende Bedrohung für jeden Internetnutzer dar. Bei Backdoor-Software handelt es sich sozusagen um unseriöse Verwandte von sogenannten **(RAT)**. Dies sind Fernwartungstematadministratoren benutzt werungsarbeiten vorzunehmen. Da ohne Zustimmung des Computeranderen Programmen versteckt Trojaner. Alle Trojaner und RATs zip. Sie bestehen aus mindestens und dem Client. Der Server muss auszuspionieren oder zu warten speicher aktiv sein. Dieser Remote-Rechner oder Host bezeichnet. Der zweite Programmteil, der Client, befindet sich auf dem Rechner, von dem aus die Kontrolle ausgeübt werden soll. Viele Trojaner und RATs enthalten außerdem noch eine dritte Komponente, die dazu dient, den Server abhängig von den jeweiligen Anforderungen zu konfigurieren. Besteht eine Verbindung zum Internet, sind prinzipiell alle ca. 65.000 Ports offen bzw. einige wenige sind konkret vergeben. Die übrigen können also in böser Absicht verwendet werden. Wenn ein Hacker einen fremden Computer ausspionieren will, ist die größte Schwierigkeit, den Trojaner auf dem Remote-Rechner unbemerkt zu installieren. Eine Möglichkeit ist z.B. der Einsatz sogenannter Exe-Binder. Im Internet stehen davon eine ganze Reihe zur Verfügung. Das sind Programme, mit denen zwei oder mehr Anwendungen zu einer einzigen ausführbaren Datei zusammengefasst werden können. Beim Start dieser Datei werden dann alle enthaltenen Programme ausgeführt. Eine weitere beliebte Art, Trojanische Pferde unter die Leute zu bringen, sind böswillige Links, Java-Applets oder ActiveX-Controls auf Web-Seiten.



**Remote Administration Tools** werkzeuge, die häufig von Sysden, um über das Internet Wardiese Backdoor-Software, die Users eingesetzt wird, häufig in wird, bezeichnet man sie auch als arbeiten nach dem gleichen Prinzwei Programmteilen, dem Server sich auf dem Computer, den es gilt, befinden und dort im Arbeiter wird normalerweise als Re-

## Cookies

Cookies sind kleine Textdateien (meist nur einige hundert Byte groß), die von einem Internetserver auf die Festplatte abgelegt werden und bei einem erneuten Besuch wieder von diesem geladen werden können. Sie dienen eigentlich als Marketing-Instrument, um kommerzielle Internetangebote individuell auf die Kunden abzustimmen. Cookies werden vom Browser immer in ein und demselben Verzeichnis abgelegt und enthalten den Domain-Namen des besuchten Servers. So können die Werbestrategen das Surfverhalten des Benutzers ausspionieren und entsprechend das Werbeangebot anpassen. Da Cookies wenig Speicherplatz belegen und keinen direkten Schaden anrichten, kann es individuell sogar Sinn machen Cookies zu akzeptieren. Ansonsten werden sie im Browser in den Sicherheitseinstellungen deaktiviert.



## Verschlüsselung

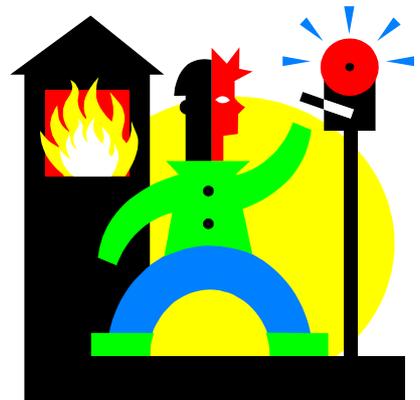
Die Kommunikation per E-mail ohne Verschlüsselungssoftware ist im Allgemeinen keine sichere Angelegenheit und sollte deshalb nicht benutzt werden, um wichtige geheime persönliche Daten



anderen mitzuteilen. Gegenwärtig werden asymmetrischen Verschlüsselungsverfahren bevorzugt. Im Gegensatz zur symmetrischen Verschlüsselung, bei der zum Verschlüsseln und zum Entschlüsseln ein und derselbe Code verwendet wird, arbeitet die asymmetrische Verschlüsselung mit einem Schlüsselpaar, nämlich dem *private* oder *secret key* und dem *public key*. Der *public key* des Mail-Empfängers wird dabei vom Absender der Mail zur Verschlüsselung verwendet und nur der "richtige" Empfänger kann die Nachricht mit seinem *private key* entschlüsseln. Dabei setzt dieses 1976 von Hellman, Diffie und unabhängig auch von Merkle erdachte asymmetrische Verfahren voraus, dass aus der Kenntnis des öffentlichen Schlüssels praktisch nicht auf den geheimen Schlüssel geschlossen werden kann. Das bekannteste dieser Verfahren ist das RSA-Verfahren (Rivest/Shamir/Adleman). Diese Verfahren sind nicht nur sicherer, sie sind auch praktikabler, da beim symmetrischen Verfahren Absender und Empfänger den Schlüssel erst vereinbaren und - geheim! - austauschen müssen.

### Firewall

Eine Firewall ist eine Schwelle meist zwischen zwei Netzen, die überwunden werden muss, um Systeme im jeweils anderen Netz zu erreichen. Es wird dafür gesorgt, dass jede Kommunikation zwischen den beiden Netzen über die Firewall geführt werden muss. Da immer wieder bislang unbekannte Sicherheitslücken in Windows entdeckt und neue Trojaner entwickelt werden, ist der einzige wirksame Schutz gegen Spionage aus dem Internet eine gut funktionierende Firewall, d.h. ein Programm zur Überwachung der Verbindungen zum Internet. Selbst größte Vorsicht und eine aktuelle Antivirensoftware garantieren nicht, dass niemand unerlaubt auf Daten zugreift. Für Rechner im lokalen Netz (LAN) übernimmt die Aufgabe der Firewall der Proxyserver.



### Proxyserver

Unter einem Proxyserver versteht man einen Computer, der für andere Rechner den Zugang zum Internet vermittelt. Der Proxyserver stellt die Verbindung zum Internet her und nur seine IP-Adresse kann deshalb identifiziert werden, der einzelne Rechner im LAN ist somit für das Internet unsichtbar. Der Proxyserver funktioniert dabei normalerweise auch als Firewall, d.h. er lässt nur autorisierte Verbindungen zu und kontrolliert alle Datenpakete, bevor sie weitergeleitet werden. Oft kommt auch gleichzeitig eine Antiviren-Software zum Einsatz, die während der Übertragung Dateien und E-mail auf schädlichen Inhalt überprüft. Trotzdem sollte man auch im LAN mit fremden Dateien vorsichtig umgehen. Ebenso wenig wird ein Proxyserver vor Gefahren durch aktive Inhalte aus dem Netz schützen, es sei denn, die Firewall wäre so konfiguriert, dass diese Inhalte automatisch blockiert werden. Dies ist aber eher selten der Fall. Daher sollten alle erwähnten Sicherheitsmaßnahmen ernst genommen und entsprechende Einstellungen vorgenommen werden.

## PROJEKTARBEIT

Projekte sind Lehrplaninhalte in nahezu allen Fachlehrplänen. Sowohl in der Informatik als auch in der Informationstechnologie nehmen sie einen besonders breiten Raum ein. Seinen Ursprung hat der Begriff im lateinischen Wort *proicere* (= *entwerfen*) und meint damit einen Entwurf, einen Plan, ein Vorhaben. Im Unterricht eingesetzt stellt es eine schüler- und handlungsorientierte und somit aktive Lernmethode dar, bei der eine Simulation von Leben, Lernen und Arbeiten im Unterricht umgesetzt werden soll. Dabei sollen die individuellen Bedürfnisse und Interessen der Schüler berücksichtigt werden und diese durch Selbstverantwortung und Selbstorganisation zu Arbeitsformen finden, wie sie später im Berufsleben als Schlüsselqualifikationen gewünscht werden. Insbesondere werden Phantasie, Kreativität, Kritikfähigkeit, Teamfähigkeit und demokratische Verhaltensweisen gefördert. Gerade die Realschule, die neben einer breit gestreuten Allgemeinbildung auch eine vorberufliche Grundlagenbildung anstrebt, muss versuchen, den Schülern die bereits angedeuteten Grundqualifikationen zu vermitteln. Sie bestehen darin, Fähigkeiten, Fertigkeiten und Verhaltensweisen zu trainieren, die ein entsprechendes Umsetzen in der späteren Berufswelt ermöglichen sollen. Als wichtigste Schlüsselqualifikationen sollen hier in alphabetischer Reihenfolge genannt werden:



- Bereitschaft zu ständigem Lernen
- Flexibilität
- Ganzheitliches Denken
- Kreativität
- Selbstkritik
- Selbstständigkeit
- Teamfähigkeit
- Urteilsfähigkeit

Projekte werden von Schülern aktiv mitgestaltet, wobei das Ziel der Projektarbeit bereits zu Beginn festgelegt wird. Im Projekt wird eine Verbindung zwischen Theorie und Praxis angestrebt, es führt zum unmittelbaren Handeln der Projektteilnehmer. Die Projektarbeit vollzieht sich stets in Gruppen, der Lehrer wirkt nur als Moderator und Berater. Der Einsatz praxisorientierter Medien ist dabei unverzichtbar. Auch die Möglichkeit, den angesprochenen Praxisbezug vor Ort kennen zu lernen, muss den Schülern eingeräumt werden. Andererseits muss man als Lehrer dahingehend motivierend wirken, dass die Schüler auch außerhalb der Unterrichtszeit selbstständig oder in der Gruppe das Projekt voranbringen.

Beispiele für mögliche Projekte (je nach Wahlpflichtfächergruppe und Klassenstufe, siehe hierzu auch die verschiedenen Fachlehrpläne):

- **Anlage von Geldvermögen**
- **Buchungskreisläufe darstellen**
- **CNC-Fertigung**

- **Entwickeln einer grafisch ansprechenden Website**  
z.B. Schul-WebSite, Präsentationen aus Kunst- oder Sportunterricht etc.
- **Entwurf, Durchführung und grafische Aufbereitung von Umfragen**
- **Erstellen einer Informationsbroschüre eines Unternehmens**

#### **Erstellen einer interaktiven Hypertext-Lerneinheit**

z.B. CAD-Konstruktionsabläufe visualisieren, Unterrichtseinheiten aus den unterschiedlichen Fachbereichen erstellen etc.

- **Finanzierungsplan für die Anschaffung einer Immobilie**
- **Logische Grundfunktionen darstellen bzw. simulieren**
- **Modellierung und praktischer Aufbau einer Datenbankanwendung**
- **Modellierung mit Hilfe eines Programmierwerkzeugs**  
z.B. Entwicklung eines einfachen Computerspiels
- **Online-Banking und E-Commerce**
- **Präsentation eines neuen Produkts bei einer Messe**
- **Programmieren eines Roboters**
- **Simulation technischer oder sozialer Vorgänge**  
z.B. Anfertigen von Explosionszeichnungen mit Hilfe eines CAD-Systems, Darstellung von einfachen Bewegungsabläufen, Modellieren im Bereich der virtuellen Realität, Simulation und Modellierung einer Ampel- oder Heizungsanlage etc.
- **Statistische Daten sammeln, strukturieren, auswerten und eine Präsentation erstellen**
- **Teilnahme an einem Wettbewerb dokumentieren**
- **Untersuchen von Verbraucherverhalten**



#### **Organisation der Projektdurchführung**

Ein Projekt kann in mehrere große Teilbereiche aufgliedert werden:

- **Erarbeiten des Projektthemas**  
In der Realschule muss die Projektinitiative natürlich vom Lehrer ausgehen, bei weniger geübten Schülergruppen wird oft auch das Projektthema weitgehend vom Lehrer vorgestellt, wobei spätestens hier die Schüler intensiv in die genaue Themenformulierung einbezogen werden.



- **Erstellen eines Ablaufplans**

Im Projektplan wird neben der genauen Festlegung der Arbeitsschritte auch die Aufteilung in Teams sowie die Aufgabenverteilung und der Terminplan festgelegt. Wichtig bei der Planung sind die themenbezogenen Dreh- und Angelpunkte, die gleichzeitig für die Schüler entscheidende Lernetappen darstellen. Hier werden Zwischenergebnisse in der eigenen Gruppe formuliert und vielleicht den anderen Gruppen vorgestellt, die weitere Vorgehensweise wird geplant und abgestimmt, evtl. notwendige Korrekturen werden besprochen und beschlossen.

- **Durchführung des Projekts**

Die Schüler erarbeiten in Gruppen mit nicht mehr als 4 Teilnehmern das ihnen zugewiesene Teilprojekt.

• **Präsentation der Projektergebnisse**

Ein Projekt ist zielgerichtet und endet mit einem Ergebnis, einem Produkt. Es werden die Gruppenarbeiten zusammengefasst. Dabei wird auch geprüft, ob die Projektziele erreicht wurden. Es werden die Stärken und Schwächen der vorliegenden Ergebnisse diskutiert und erörtert, ob gegebenenfalls nachgebessert werden muss. Das Gesamtergebnis kann zunächst allen Teams präsentiert werden, aber auch anderen Schülergruppen, Lehrern, der Öffentlichkeit etc.

• **Bewertung des Projekts (Evaluation)**

Jedes Projekt muss beurteilt werden. Das geschieht grundsätzlich durch den (oder die) Lehrer, es können aber auch die Schüler mit einbezogen werden. Dabei sollten die Bewertungskriterien klar festgelegt sein.



Kennzeichen solcher Unterrichtsprojekte ist ganzheitlich orientiertes Denken der Schüler. Die Gestaltung erfolgt problemorientiert, praxisbezogen und fächerübergreifend. Es handelt sich bei jedem Projekt um eine Simulation der Wirklichkeit, bei der die Schüler auch Fehler machen dürfen, ohne bei Verfehlung des Ziels mit schwerwiegenden Konsequenzen rechnen zu müssen. Sie lernen auch die Schwierigkeit kennen, in einem vorgegebenen Zeitrahmen zu einem brauchbaren Ergebnis zu kommen.

Da Projektarbeit in verschiedenen Jahrgangstufen und unterschiedlichen Fachbereichen erfolgt, arbeiten die Schüler in wechselnden Gruppen und erproben ihre Teamfähigkeit, lernen Arbeitsabläufe zu organisieren, beschaffen sich selbst fehlende Informationen aus unterschiedlichen Medien, setzen sich kritisch damit auseinander und bewerten diese, strukturieren die beschafften Informationen und erproben sie in der Praxis. Für die Präsentation und Dokumentation müssen wir den Schülern geeignete Werkzeuge zur Verfügung stellen und sie mit deren Möglichkeiten vertraut machen.

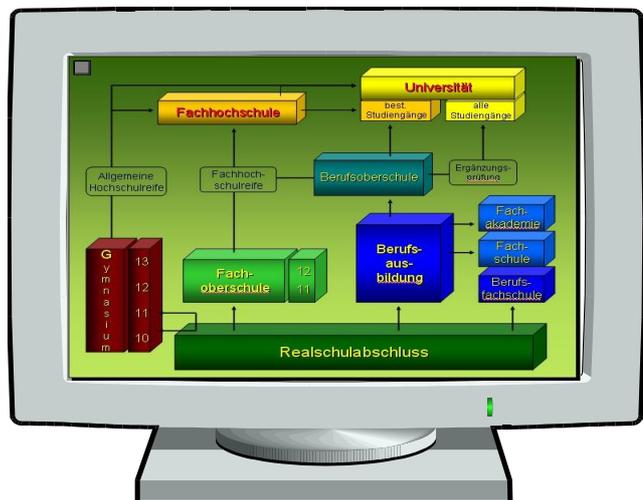
Layoutgrundlagen

**DTP**

Texte sind bei der Erstellung unstrukturiert. Bei ihrer Gestaltung werden sie in eine bestimmte Form gebracht.

Dies erfolgt meist in folgenden Schritten:

- **Texterfassung**  
(Textverarbeitungsprogramm, Scannen + Texterkennung ...)
- **Bilderfassung** (Foto + Scannen, elektronisches Foto...)
- **Gestaltung**
- **Ausgabe**



Durch **WYSIWYG** (What You See Is What You Get) kann man bereits am Bildschirm das Ergebnis so sehen, wie es später ausgedruckt erscheinen wird. Durch neue Medien (Video, Internet usw.) beschränkt sich die Gestaltung nicht nur auf Druckerzeugnisse. Gleichzeitig können neue, dynamische und interaktive Elemente eingebunden werden.

Die **Layout**-Gestaltung ist kein neues Wissen, das es erst seit der Erfindung des DTP gibt, sondern altes Drucker - Know - How. Eine gut gestaltete Seite (auch eine Webpage) zeichnet sich dadurch aus, dass sie vielen der klassischen Satzregeln entspricht.



### **Hauptaufgabe des Designs/Layouts:**

- **Aufmerksamkeit erregen**
- **den Blick auf den Inhalt ziehen**
- **eine „Botschaft“ übermitteln**

Das Layout ist Teil dieser Botschaft und so kann z.B. Seriosität oder Innovationsfreudigkeit angezeigt werden. Durch Effekte (z.B. Kontraste...) kann das Design interessant gemacht werden. Hier gilt allerdings die Regel „*Weniger ist mehr*“. Leicht wirkt das Gestaltete konfus oder abstoßend. Grundgebot ist die *Kontinuität der Ausführung* - ob seriös, modern, verspielt, kitschig oder chaotisch: einheitliches Design muss durchgehalten werden.

Ob man die Umsetzung direkt in einem Textverarbeitungsprogramm vornimmt oder ein für diese Zwecke noch besser geeignetes rahmenorientiertes Desktop-Publishing-Programm (wie z.B. MS Publisher) zur Verfügung hat ist zunächst zweitrangig, trotzdem sollen hier einige Hinweise dazu gegeben werden:

In die leere Seite werden alle Objekte (hier Rahmenelemente) eingegeben, die später Daten enthalten sollen:

- |                    |   |                                   |
|--------------------|---|-----------------------------------|
| • Fließtext        | ⇒ | Textbox                           |
| • gestalteter Text | ⇒ | WordArtBox (Microsoft-spezifisch) |
| • Bild             | ⇒ | BildBox                           |
| • Grafik           | ⇒ | Clip-Art-Box usw.                 |

Alle Objekte sind frei verschiebbar und müssen zur Bearbeitung ausgewählt werden. Die unterschiedlichen Eigenschaften der Objekte werden entweder über Menüs bzw. Icons ausgewählt und können dementsprechend geändert werden.

Auf der Arbeitsfläche ist die Seite sichtbar. Zusätzlich können Objekte auf dem Arbeitsplatz d.h. rund um die Seite abgelegt werden. Weitere Einzelheiten sind weiterführender Literatur vorbehalten bzw. können über die Online-Hilfe abgerufen werden.

## Erstellung elektronischer Bilder (Scanner, E-Photo, Grafikprogramm)

Bilder sind für aussagekräftige Publikationen unbedingt notwendig. „*Ein Bild sagt mehr als 1000 Worte*“. Um Bilder im Electronic-Publishing verwenden zu können, müssen sie 'in den Computer gebracht' werden. Hierzu gibt es verschieden Möglichkeiten:

### Scanner

Mit einem Lichtstrahl wird eine (herkömmliche) Fotovorlage abgetastet und in digitale Informationen umgewandelt. Für die Software besteht ein Bild aus zahlreichen Punkten, die sich entsprechend verarbeiten lassen. Probleme bereiten Bilder, die bereits einmal gedruckt waren, da sie schon aus Einzelpunkten (**Raster**) bestehen. Im Druck erscheint häufig ein Moiréeffekt, d.h. Kreise und andere Muster.

Grundregel für die Scannereinstellung:

Internet            75 – 100 dpi  
Druck                ½ Druckerauflösung (300 – 600) (Ausnahme: Strichzeichnungen - \* 2)



### Elektronische Kamera

Bilder werden direkt digital erzeugt und elektronisch gespeichert. Sie liegen damit bereits in Pixel-Form vor. Die Qualität hängt von der Auflösung (Pixel pro Fläche) vor.

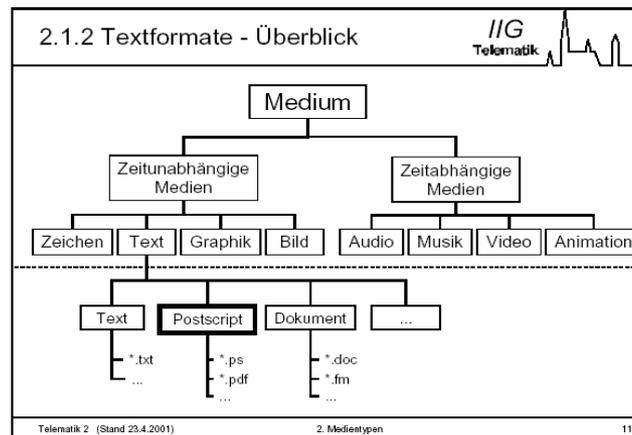
### Grafikprogramme

Ein Bild oder eine Zeichnung wird mit Hilfe des Computers „gemalt“ (oder verändert – siehe 1.4). Hierbei gibt es 2 Möglichkeiten: Bei **Bitmaps** wird an jeder Stelle des Bildes ein Punkt erzeugt (oder verändert). (Fotos lassen sich nur auf diesem Weg bearbeiten.) **Vektorbilder** sind dagegen über eine mathematische Beschreibung von Linien, Umrissen und Füllungen festgelegt. Sie können „glatt“ vergrößert und verkleinert werden, ohne ein grobes Sägezahnmuster („Treppchen“) zu erhalten.

Im Computer werden Grafiken in verschiedenen Grafikformaten gespeichert. Man erkennt sie am einfachsten über die Dateiendungen. Hier einige Beispiele:

<b>BMP</b>	Windows Bitmap-Format, mind. 75 dpi, normal 300-600 dpi
<b>WMF</b>	Windows Metafile, Austauschformat mit relativ geringer Qualität
<b>GIF</b>	Graphic interchange format, hohe Komprimierung (= kleiner Speicherbedarf), geringere Qualität (max. 256 Farben)
<b>JPEG</b>	Joint Picture Experts Group, internationaler ISO-Standard, v.a. für Fotos und Halbtondarstellungen gut geeignet;
<b>TIFF/TIF</b>	Tagged image file Format, sehr hohe Druckqualität durch (leider unterschiedlich) standardisierte Beschreibungssprache

## Textformate



### ASCII

Der mit 7 bit zu beschreibende *American Standard Code for Information Interchange* (ASCII) ist der am meisten verbreitete Kommunikationscode nicht nur innerhalb der Vereinigten Staaten, sondern weltweit.

### RTF

Fast jedes Text-Programm besitzt eine Möglichkeit, Dateien als RTF zu speichern und zu öffnen. RTF (Rich Text Format) unterstützt grundlegende Formatierungen wie Zeilenabstände, Tabulatoren, Schriftarten.

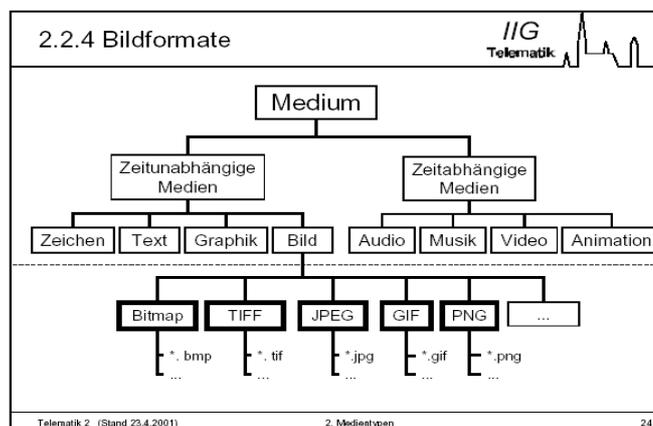
### HTML

Die *HyperText Markup Language* (HTML) ist der Internet-Standard, der auf Basis des ASCII-Zeichensatzes den "Internet Hypertext" definiert. Als Hypertext bezeichnet man durch Querverweise verbundene Dokumente. HTML enthält generische Kodierung zur Beschreibung von Dokument- und Verweiselementen, ähnlich PostScript. HTML ist von der generischen und international standardisierten Seitenbeschreibung SGML abgeleitet worden

### PostScript

PostScript wurde von der Firma Adobe ca. 1985 konzipiert und ist in erster Linie eine ASCII-basierte, komplexe Seitenbeschreibungssprache, die hauptsächlich zur Ansteuerung von Laserdruckern benutzt wird.

## Bildformate



## Bitmapgrafik

Pixelorientierte Bilder (Bitmaps) setzen sich aus einzelnen Punkten (Pixeln) zusammen. Diesen Pixeln werden die Farben zugeordnet. Eine Linie besteht also durch eine Aneinanderreihung mehrerer Bildpunkte. Werden diese Grafiken vergrößert (gezoomt), entstehen Rastereffekte, die die Bildschärfe stark beeinträchtigen. Der punktförmige Aufbau entspricht gleichermaßen der Bildentstehung am Monitor oder Drucker wie der im menschlichen Auge.

- Die Grafik besteht logisch aus einzelnen Pixel.
- Jedes Pixel wird einzeln betrachtet und beschrieben.
- Es besteht keine Abhängigkeit der Punkte zueinander.
- Einzelne Bestandteile der Grafik lassen sich nur schwer wieder aus der Grafik entfernen.

## Vektorgrafik

Vektorgrafiken bieten sich für geometrisch strukturierte Vorlagen an. Mit Koordinationsangaben wird die Platzierung von Objekten wie Linien, Kreisen, Ecken, Kurven etc. beschrieben. Vektorgrafiken können beliebig gezoomt werden, ohne dass Rastereffekte entstehen.

- Nicht Pixel sind die Objekte, sondern Linien, Rechtecke, Ellipsen usw.
- Jedes Objekt wird durch mathematische Ausdrücke (Vektoren) definiert.
- Von einer Linie z.B. werden als Informationen nur Anfangspunkt und die zugewiesenen Attribute (Länge, Richtung, Farbe, Strichstärke usw.) gespeichert.
- Das Aussehen der Objekte ergibt sich aus vergleichsweise wenigen Daten, wie z.B. Punktkoordinaten, Ellipsenradien oder ähnlichem.
- Eine mit Vektoren definierte Grafik kann beliebig vergrößert oder verkleinert werden, ohne merkliche Verschlechterung der Ausgabequalität.
- Jedes Objekt kann als ganzes jederzeit wieder aus der Grafik entnommen werden.

## Farbtiefe

Die Farbtiefe beschreibt die Anzahl der maximal gleichzeitig nutzbaren Farben, nicht zu verwechseln mit der Anzahl der genutzten Farben! Die Farbtiefe ergibt sich aus der Anzahl der Bits, die pro Pixel zur Verfügung stehen, also 2 hoch Anzahl der Bits.

Gebräuchliche Farbtiefen sind:

1 Bit = Schwarzweiß (Bitmap) 2 Farben

4 Bit = 16 Farben

8 Bit = 256 Grautöne oder indizierte Farben

16 Bit = 65.536 Farben (High Color)

24 Bit = 16.777.216 Farben (True Color)

32 Bit = *CMYK-Modell (Cyan, Magenta, Yellow, Black)*

### 256 Farben:

Wenn jedes Pixel genau 1 Byte (8 Bit) belegt, lassen sich 256 verschiedene Farben darstellen:

1 Pixel = 1 Byte = 8 Bit →  $2^8 = 256$  Farben

Die Farbpalette kann in vielen Programmen individuell festgelegt werden.

**High Color:**

Die 16 Bit werden auf die Farbanteile des RGB-Modells verteilt: 6 Bit rot, 5 Bit grün und blau. Für rot 1 Bit mehr, weil das Auge mehr Rottöne unterscheiden kann.

1 Pixel = 2 Byte = 16 Bit →  $2^{16} = 65536$  Farben

**True Color:**

Für jeden Farbanteil der RGB-Palette stehen 8 Bit zur Verfügung, obwohl das Auge so viele Farbtöne gar nicht wahrnehmen kann.

1 Pixel = 3 Byte = 24 Bit →  $2^{24} = 16,78$  Mio Farben

**Interlaced**

Beim Interlaced -Verfahren werden die Bilder in einzelnen Schärfestufen aufgebaut, so dass bereits nach der halben Übertragungszeit ein Bild erkennbar ist. Der Bildaufbau erfolgt vom verschwommenen ins scharfe.

**Transparenz**

Eine beliebige Farbe kann transparent dargestellt werden, um z. B. transparente Hintergründe zu erzeugen. Drei Grafikformate (GIF, JPG, PNG) werden von Webbrowsern im Moment standardmäßig unterstützt, da sie sich gut für die Übertragung im Internet eignen. Bei der Entscheidung über das zu verwendende Format spielen folgende Gesichtspunkte eine Rolle:

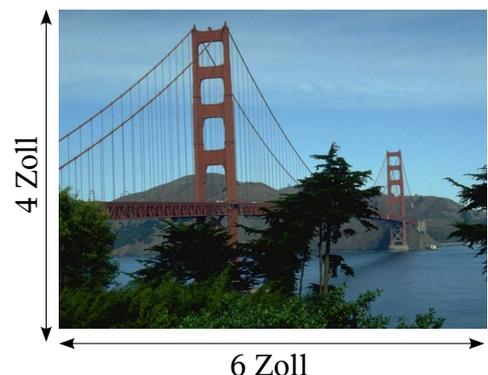
- *der zu erstellende Bildtyp*
- *die Dateigröße (Grafiken sollten möglichst wenig Speicherplatz verbrauchen, damit sie schnell über das Netz geladen werden können)*
- *die gewünschte Bildqualität (möglichst gut)*
- *die Geschwindigkeit beim Anzeigen auf dem Bildschirm (möglichst schnell)*

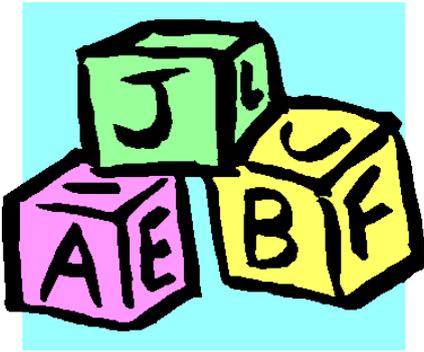
Allgemein gilt das GIF-Format als beste Lösung, wenn es um Strichgrafiken und Grafiken mit wenigen Farben oder scharfen Rändern geht. JPEG ist das bevorzugte Format zum speichern von Bildern mit zahlreichen Farben, beispielsweise Fotografien oder eingescannte Bilder. PNG ist ein relativ neues Format, das die Vorteile der beiden anderen Formate bis auf einige Details (Animationen) vereinigt.

**.BMP (Bitmap)**

BMP ist das unter dem Betriebssystem Windows 3.0 entstandene und durch die weiteren Versionen dieser Betriebssystemfamilie häufig verwendete Bildformat. Farbtiefen von 1 bis 24 Bit sind einstellbar. Da es sich um ein 'Pixel-Format' handelt, sind die zugehörigen Dateien immens groß. Ein postkartengroßes Bild (10 x 15 cm entspricht etwa 4 x 6 Zoll) hat bei einer Auflösung von 300 dpi (dots per inch) und einer Farbtiefe von 24 Bit (True Color) eine Dateigröße von 6,2 MByte:

1200 Pixel x 1800 Pixel x 3 Byte = 6 480 000 Byte



**.GIF (Graphics Interchange Format)**

GIF wurde als plattformübergreifender Standard entwickelt und wird von allen grafischen Internet-Browsern unterstützt. GIF unterstützt Farbtiefen bis 8 Bit (256 Farben möglich) und ermöglicht das Speichern von angepassten Paletten gemeinsam mit dem Bild. GIF bietet einige erweiterte Grafikoptionen, darunter transparente Hintergründe, Bild-Interlacing und Animation. GIF bietet verlustfreie Komprimierung, d.h. beim Konvertieren in das GIF-Format werden alle vorhandenen Dateiinformationen mit dem Bild gespeichert, und die GIF-Datei entspricht exakt der ursprünglich von Ihnen erstellten Grafik. Da nur eine begrenzte

Dekomprimierung erforderlich ist, erscheinen die Grafiken im GIF-Format ziemlich schnell auf dem Bildschirm (zum Vergleich: Bitmap-Dateien im BMP-Format sind bei gleichem Inhalt durchschnittlich zehnmal bis dreißig mal so umfangreich wie GIF-Dateien)

**.JPEG (Joint Photographic Experts Group)**

Das JPEG-Grafikformat komprimiert ebenfalls sehr gut und hat gegenüber dem GIF-Format den Vorteil, dass es pro Bild 16,7 Millionen Farben speichert. Der Nachteil ist, dass JPEG mit Verlust komprimiert, d.h. je höher der Komprimierungsfaktor, desto schlechter die Qualität der Grafik. Aufgrund seiner Charakteristik eignet sich das JPEG-Format vor allem zum WWW-gerechten Abspeichern von eingescannten Fotos, aber auch für andere Grafiken, in denen sehr feine Farbverläufe vorkommen. Beim JPEG-Format können Sie bei besseren Grafikprogrammen den Komprimierungsfaktor selbst bestimmen.

**.PNG (Portable Network Graphic)**

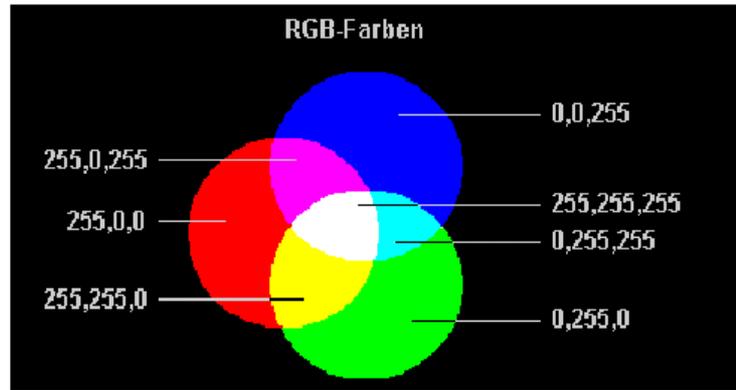
PNG bedeutet Portable Network Graphic (ausgesprochen: PING). Es handelt sich um ein Grafikformat, das eigens für den Einsatz im WWW konzipiert wurde. PNG soll die Vorteile von GIF mit denen des JPEG-Formats vereinen, wird aber von älteren Browsern nicht unterstützt.

<i>Vergleich der Grafikformate</i>		
<b>*.jpg</b>	<b>*.gif</b>	<b>*.png</b>
verlusthaltige Kompression	verlustlose Kompression (nur bis 256 Farben)	verlustlose Kompression
16,7 Millionen Farben	256 Farben	bis 4,27 Milliarden Farben
keine Transparenz	keine transparente Farben	Transparenz
interlacing (progressive-JPEG)	interlacing	Interlacing
keine Animationen	Animationen	keine Animationen
nur zum Export geeignet, da bei jedem Speichern neu komprimiert wird (Verlust an Qualität)		PNG-Dateien sind größer als JPG- bzw GIF-Dateien

**RGB (Rot, Grün, Blau)**

Eine Farbe wird durch die drei Grundfarben Rot, Grün und Blau dargestellt. Jeder dieser Farben wird dabei ein Wert zwischen 0 und 255 zugeordnet. Der Wert 0 besagt, dass kein Anteil an der betreffenden Grundfarbe vorhanden ist und Wert 255 steht für den Maximal-Anteil der Farbe.

Rot, Grün und Blau sind Farben, die sich nicht durch Mischen anderer Farben erzeugen lassen. Es handelt sich um additive Primärfarben. Durch Mischen zweier additiver Primärfarben entsteht eine Sekundärfarbe, wobei auch Grauwerte und Schwarz erzeugt werden. Mit dieser additiven Farbmischung lassen sich 16,7 Mio. Farben darstellen.



### Komprimierungsverfahren

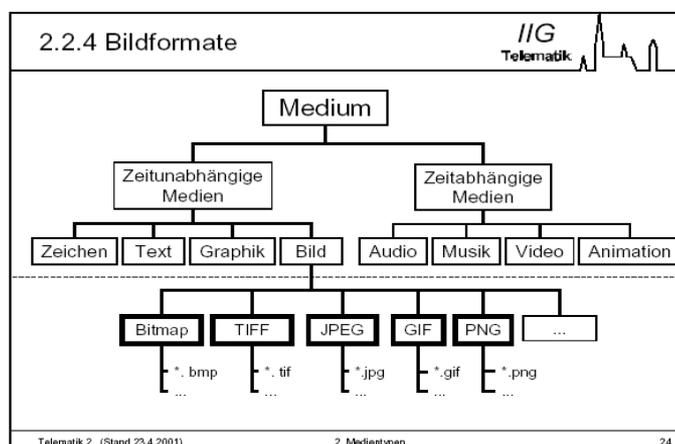
Die drei wichtigsten Komprimierungsverfahren sind die Lauflängen-Codierung, die LZW-Codierung und die JPEG-Kompression.

Bei der **Lauflängen-Codierung (RLE = Run Length Encoding)** wird neben dem Farbwert eines Pixels noch die Anzahl von nachfolgenden identischen Pixeln gespeichert. Da dieses Verfahren zeilenweise arbeitet wird eine Datei mit vielen horizontalen Linien stärker komprimiert als Dateien mit vertikalen Linien. Es handelt sich um ein verlustfreies Komprimierungsverfahren.

Das nach seinen Erfindern benannte **LZW-Verfahren (Lempel, Ziv, Welch)** durchsucht eine Datei nach gleichen Musterfolgen. Diese häufig auftretenden Muster erhalten einen Code, der in der komprimierten Datei die Pixelfolge repräsentiert. Die Art der Codierung erfolgt nach einer festliegenden Code-Tabelle, so dass diese nicht mit gespeichert werden muss. Die Kompression ist hier ebenfalls verlustfrei.

Das komplexere **JPEG-Komprimierungsverfahren** arbeitet nicht verlustfrei. Das Bild wird in lauter 8 x 8 = 64 Pixel große Blöcke zerlegt. Da jedes Pixel von jedem Farbwert einen bestimmten Farbanteil besitzt wird dieses Verfahren für jeden Farbwert gesondert durchgeführt. Für jeden Farbwert existieren 64 Grundmuster (Frequenzen). Für jeden 8 x 8 – Block wird durch die sog. diskrete Kosinus-Transformation durch Überlagerung von mehreren Grundmustern. Die eigentliche Kompression erfolgt erst in einem weiteren Schritt, der Quantisierung: da die Farbanteile der höher-frequenten (d.h. mit stark variierenden) Verläufe sind erfahrungsgemäß fast gleich groß sind, werden sie auf einen einheitlichen Mittelwert gesetzt. Diese gleichen Zahlenwerte lassen sich dann komprimieren. Die Mittelung verfälscht die Farbwerte im Block geringfügig, was das menschliche Auge kaum wahrnimmt. Je höher der Kompressionsgrad desto mehr Frequenzen werden gleichgesetzt.

### Audioformate



### MP3

Die MPEG 2.5 Audio Layer III (MP3) -Technologie ermöglicht es, digitale Informationen, zum Beispiel Musik, auf ein Zwölftel der ursprünglichen Datenmenge zu komprimieren, ohne dass dabei die Qualität des Musikstücks merklich abnimmt. Ursprünglich wurde das Format von der Fraunhofer-Gesellschaft (FHG), jedoch als Kompressionsverfahren für Audio-Kanäle auf Video-CDs entwickelt.

Der Trick: Die Entwickler von MP3 lassen bei der Umwandlung der Musik einfach alle Klangteile weg, die das menschliche Ohr sowieso nicht wahrnehmen kann, zum Beispiel ein Flüstern, das von einer Explosion überdeckt wird.

### Midi



Midi bezeichnet ursprünglich den Standard für ein Interface um Daten zwischen einem Computer und elektronischen Musikinstrumenten auszutauschen.



Applause.wav

Die Klänge werden von der Soundkarte selbst erzeugt, ähnlich einer Notenschrift werden die einzelnen Klänge durch ihre Dauer und Tonhöhe (sowie je nach Instrument weiteren Eigenschaften) beschrieben. Der Vorteil liegt in sehr kleinen Dateien und geringem Rechenzeitverbrauch, der größte Nachteil besteht in der Komplexität des Formates, von einer Midi-Datei die für einen mehrere 1000 € teuren Profi-Synthesizer mit Hunderten von Klängen verfasst wurde, kann man nicht erwarten, dass sie auf einer Soundkarte für einige 100 € auch nur annähernd ähnlich klingt.

### Wave-Dateien

Wavedateien können beispielsweise in dem Format WAV vorliegen und bestehen aus digitalen Audioaufzeichnungen. Damit lässt sich neben Musik auch Sprache, ähnlich einem Tonbandgerät aufzeichnen. Die üblichsten Qualitäten sind hierbei 8 und 16 Bit Aufzeichnung in 11, 22 oder 44 kHz. Als Richtwerte gelten: 8bit/11kHz ist etwa Telefonqualität, 16Bit/22kHz etwa Radio- und 16Bit/44 kHz etwa CD-Qualität.

### Video-Formate

Bei der Videobearbeitung fallen sehr große Datenmengen an. Also stellt sich die Frage, ob man nicht mit Blick auf das Ergebnis die anfallenden Datenmengen durch eine geringere Auflösung reduziert. Für VHS wird eine Auflösung von 320 x 200 Bildpunkten benötigt. S-VHS und DV arbeitet mit 720 x 576 Bildpunkten. Die volle PAL-Auflösung beträgt 768 x 576 Bildpunkte, bei NTSC werden 640 x 480 Bildpunkte verwendet.

Da es sich bei NTSC- und PAL-Videosignalen um analoge Signale handelt, Computer-informationen jedoch digital anzeigen, müssen sie digitalisiert oder "gesampelt" werden, bevor sie mit einem Computer verarbeitet werden können. Der Vorgang der Digitalisierung wird häufig auch als Aufnehmen bezeichnet. Video-Digitalisierungskarten (häufig auch als Frame-Grabber bezeichnet) werden zum Aufnehmen und Umwandeln eines analogen Signals in ein Computer-Signal benutzt. Das digitale Aufnehmen eines Videosignals benötigt eine erhebliche Menge an Speicherplatz, da die Farb- und Helligkeitsinformation jedes Pixels in jedem Frame gespeichert werden muss. Ein NTSC-Bild in der vollen Bildschirmgröße eines 13" Computermonitors misst 640 x 480 Pixel. Somit enthält jedes Frame eines Vollbild-Videos 307.000 Pixel (640 x 480 Pixel). Um ein solches Vollbild digital in 24-Bit Farbtiefe (8-Bit für jede der RGB-Komponenten) darzustellen, muss jeder

Pixel 24-Bit an Farbinformation darstellen. Da 24-Bit 3 Bytes entspricht benötigt jeder digitalisierte Frame bei 307.200 Pixeln pro Frame 921.600 Bytes. Bei einer Framerate von 30 fps werden etwa 27 Megabyte zum Speichern von einer Sekunde an digitalisiertem NTSC-Video benötigt! Führt man dieselbe Berechnung für ein PAL-Bild (768 x 576 Pixel) durch, enthält jeder Frame eines Vollbild-Videos 442.367 Pixel. Bei 24-Bit Farbtiefe und einer Framerate von 25 fps benötigt man einen Speicherplatz von 32 MB für eine Sekunde an digitalisiertem Film. Der Umgang mit solchen Speichermengen ist für die meisten Computeranwender nicht realisierbar.

Ein noch größeres Hindernis ist die Rechenleistung, die zum Abspielen der gespeicherten Informationen mit ausreichenden Frameraten benötigt wird. Damit Videos an einem Desktop Computer verarbeitet werden können, sind Fortschritte bei der Entwicklung von Datenkomprimierungstechnologien und Kompromisse in Bezug auf Framegröße, Farbtiefe und Bildauflösung notwendig. Die meisten Fortschritte wurden bisher in der Art erzielt, wie Daten komprimiert werden. "Video-Grabber" dienen zur Digitalisierung von Standbildern und Videosequenzen. Die mittels Video-Grabber aufbereiteten (sprich: digitalisierten) Informationen werden sodann als herkömmliche Bild- oder Video-Datei auf die Festplatte des PCs gespeichert und können sodann auf gewohnte Weise mit dem Personalcomputer weiterbearbeitet werden.



### Video-Komprimierung

Die digitale Darstellung eines Studio-TV-Signals erfordert eine Nettodatenrate von 166 Mbit/s. Keine der bereitstehenden digitalen Übertragungsmedien im Computerbereich kann eine solche Datenrate bieten. Mit den hier behandelten Standards können jedoch auch Video- und Audio-Signale auf bereits bestehende Medien übertragen und gespeichert werden, da sie die Menge der Daten durch verschiedenste Techniken reduzieren bzw. komprimieren. Komprimierung ist der Vorgang des Entfernens oder Umstrukturierens von Daten, um die Größe einer Datei zu verringern. Digitale Videodateien sind sehr umfangreich und benötigen hohe Datenübertragungsraten zum Aufnehmen und Abspielen. Beim Kompilieren einer "Video für Windows"- oder QuickTime-Datei werden die Daten komprimiert, um die Dateigröße zu verringern und die Wiedergabe des Films zu vereinfachen. Die Daten werden bei der Wiedergabe des Films dekomprimiert. Für das Komprimieren von Video für Windows oder QuickTime Filmen sind einige Komprimierungs-/ Dekomprimierungs-Algorithmen (Codecs) verfügbar. Ein Codec ist ein Algorithmus zur Kompression (COmpressor) Dekompression (DECompressor) von Multimediadateien - also z.B. von Video für Windows- und QuickTime-Filmen. Codecs können auf Software oder auf Hardware basieren, wobei Hardware-Komprimierung sehr viel schneller und effektiver ist als Software-Komprimierung. Der gewählte Algorithmus beeinflusst die visuelle Qualität des Films und die Geschwindigkeit, mit der er am Bildschirm des Computers oder an einem NTSC- bzw. PAL-Bildschirm wiedergegeben werden kann. Im allgemeinen können 24-Bit Video-Vollbilder nur mit Hardware-Komprimierung und Dekomprimierung in Echtzeit abgespielt werden (d. h. bei normaler Wiedergabegeschwindigkeit). Videos für das Abspielen von einer CD werden normalerweise per Software komprimiert, um dem, der über einen CD-Player verfügt, das Abspielen der Filme ohne spezielle Hardware zu ermöglichen.

Als Video-Standard gilt das **AVI**-Format. AVI steht für „Audio Video Interleave“.

- Es ist eines der ältesten Audio-/Videoformate von Microsoft.
- Es wurde für die Wiedergabe von kurzen Videoclips entwickelt.
- Audio- und Videodaten werden ineinander verzahnt („interleaved“) abgespeichert.
- Es kommt die so genannten Keyframe-Technik zum Einsatz:
  - Jedes 12. – 17. Bild wird als Vollbild abgespeichert

- Für die dazwischenliegenden Frames werden die Unterschiede zum jeweils vorhergehenden Bild angegeben



## Das MPEG-Format

Da die digitale Darstellung eines Studio-TV-Signals eine Nettodatenrate von 166 Mbit/s erfordert und keines der bereitstehenden digitalen Medien eine solche Datenrate bieten kann, war Kompression oberstes Gebot. Video und Audio sind zeitabhängige Medien, was es ermöglicht z.B. statt der einzelnen Bilder nur die Unterschiede zwischen aufeinander folgenden Bildern abzuspeichern. Außerdem kann eine Verringerung des Informationsgehaltes pro Bild durch die schnelle Bildfolge (also die teilweise Wiederholung von Bildteilen) wettgemacht werden. Deshalb besteht keine Notwendigkeit, alle digitalen Informationen zu codieren.

MPEG wurde durch die Wahl der Codierung auf asymmetrische Anwendungen (Fernseh- Technik, Multimedia-Mail, etc.) zugeschnitten, das heißt die Herstellung des digitalen Datenstroms erfordert wesentlich mehr Ressourcen als dessen Konsumierung. MPEG ist nicht die optimale Videocodierung, allerdings ist sie der einzig internationale Konsens in diesem Bereich und findet, auch dank seiner Verwandtschaft mit den Telekommunikations-Standards (z.B S-VHS) sowie seiner Kompatibilität zu allen CD-ROM-Formaten weite Verbreitung. Prinzipiell kann derzeit zwischen vier Standards unterschieden werden: MPEG-1, MPEG-2, MPEG-4 und MPEG-7.

**MPEG-1** wurde 1993 mit dem Ziel verabschiedet, für Medien mit geringer Bandbreite (1 MBit/s bis 1,5 Mbit/s) die Übertragung von bewegten Bildern mit zugehörigem Audiosignal bei akzeptabler Bildwiederholffrequenz und möglichst guter Bildqualität zu erreichen. Das Entwurfsziel bei MPEG-1 ist der wahlfreie Zugriff auf eine Sequenz innerhalb einer halben Sekunde, ohne das dabei merkliche Qualitätsverluste auftreten dürfen. Für die meisten Heimanwendungen (Digitalisierung von Urlaubsvideos) sowie für den Businessbereich (Image-Videos, Dokumentation) ist die Qualität von MPEG-1 ausreichend.

**MPEG-2** gibt es seit 1995 und stimmt in seiner Grundstruktur mit dem MPEG-1-Format überein. Es erlaubt Datenraten bis zu 100 MBit/s und kommt bei digitalem Fernsehen (DF1), Videofilmen auf DVD-ROM und in professionellen Videostudios zum Einsatz. MPEG-2 ist in Auflösung und Datenrate über einen weiten Bereich skalierbar. Aufgrund seiner hohen Datenrate gegenüber MPEG-1 und dementsprechend hohem Speicherplatzbedarf eignet sich MPEG-2 derzeit im Heimbereich nur zur Wiedergabe. Ab einer Datenrate von zirka 4 MBit/s ist die erzielbare Videoqualität deutlich besser als bei MPEG-1.

**MPEG-4** zählt zu den neuesten Videoformaten und verfolgt das Ziel, eine möglichst hohe Videoqualität bei extrem niedrigen Datenraten im Bereich zwischen 10 KBit/s und 1 MBit/s zu erreichen. Weiterhin steht die Forderung nach Robustheit und fehlerfreier Datenübertragung, was besonders im Mobilfunk eine Rolle spielt. Eine völlige Neuerung an MPEG-4 ist die Gliederung des Bildinhalts in eigenständige Objekte, um sie gezielt anzusprechen oder weiter zu verarbeiten. MPEG-4 kommt beispielsweise bei der Videoübertragung übers Internet zum Einsatz. Einige Hersteller planen, in Zukunft bewegte Bilder per Handy zu übertragen. MPEG-4 soll dann die Basis für den Datentransfer bilden.

**MPEG-7** ist das jüngste Projekt aus der MPEG-Familie. Es stellt einen Standard zur Beschreibung von Multimedia-Daten dar, der unabhängig von den anderen MPEG-Standards eingesetzt werden kann.

## Streaming

Durch "Streaming" können Audio- und Video-Dateien bereits während der Übertragung (z.B. im Internet) angehört bzw. angeschaut werden - vorausgesetzt die Bitrate ist nicht höher als die Geschwindigkeit des (Internet-)Anschlusses. Beim Streaming muss also nicht wie beim üblichen "Store-and-forward-Prinzip" abgewartet werden, bis eine Media-Datei (Audio oder Video) komplett übertragen ist, bevor der Inhalt angehört bzw. betrachtet werden kann. Man spricht daher auch von Echtzeitübertragungen, da Daten ohne größere Zeitverschiebung vom Client wiedergegeben werden.

Um beispielsweise eine CD-ähnliche Audioqualität zu erreichen, wird eine Bitrate von 128 kb/s benötigt. Das bedeutet, dass zwei ISDN-Leitungen mit je 64kbit/s und eine "staufreie" Internetverbindung nötig wären, um MP3's direkt aus dem Internet abspielen zu können. Wenn diese Voraussetzungen nicht erfüllt werden, dann wird eine Kombination aus Streaming und Herunterladen eingesetzt: Bevor das eigentliche Abspielen beginnt, wird ein Teil der Audio- bzw. Video-Daten herunter geladen und zwischengespeichert (buffering). Dieser Teil liegt dann lokal auf dem Computer des Konsumenten und kann somit problemlos abgespielt werden. Während dieser zwischengespeicherte Teil angehört bzw. betrachtet wird, werden automatisch im Hintergrund bereits die nächsten Teilstücke herunter geladen. Dadurch kann trotz langsamerer oder manchmal verzögerter Verbindung oft ein kontinuierliches Abspielen erreicht werden. Während der Übertragung kann jederzeit zurück- und vorgespult oder die Wiedergabe gestoppt werden. Diese Technik wurde für verschiedene Übertragungsgeschwindigkeiten optimiert - also z.B. für Modem, ISDN und Netzwerk. Streamingformate: Für das Audio- und Video-Streaming gibt es unterschiedliche Datei- und Übertragungsformate (MP3, Real-Audio, Real-Video, Quicktime, WMA, ...). Die jeweilige Abspielsoftware bzw. das passende Browser-PlugIn kann bei den Herstellern in der Regel kostenlos erhalten werden. Aktuelle häufig genutzte Formate sind: VxTreme Web Theater, RealPlayer, WM (Windows Media Audio), Adobe QuickTime 3/4.

