

# 1 Einführung – Computersysteme und Informatik



- Die aufgeführten Begriffe erklären und in den richtigen Kontext einordnen können.
- Den Aufbau und die Funktionsweise eines Computersystems darstellen können.
- Gegebenen Aufgabenstellungen die jeweils geeigneten Peripheriegeräte zuordnen können.
- Die Wissenschaftsdisziplin Informatik mit ihren Definitionen, ihren Positionen und ihren Teildisziplinen erläutern können.
- Anhand von Beispielen die Arbeitsweise der Zentraleinheit zeigen können.
- Anhand von Beispielen Übertragungszeiten bei vernetzten Computersystemen berechnen können.
- Die bei der Installation und Benutzung von Software ablaufenden Vorgänge anhand von Beispielen in zeitlich richtiger Reihenfolge durchführen können.
- Software installieren und benutzen können.



**Prof. Dr. Johann von Neumann**

\* 1903 in Budapest, Ungarn  
 † 1957 in Washington D.C., USA;  
 Wegbereiter der amerikanischen Computerentwicklung, Hauptidee: Gemeinsamer Speicher für Programme und Daten, wodurch sich das Programm selbst verändern kann (1946); so aufbaute Computer mit einem Prozessor bezeichnet man heute als von-Neumann-Computer bzw. von-Neumann-Architektur; wesentliche Beiträge zu den Computern Harvard Mark I (ASCC) und ENIAC; Studium der Mathematik in Budapest, 1928 Habilitation in Berlin, danach in Berlin und Hamburg tätig, ab 1929 in Princeton (N.J.), seit 1933 dort Professor für Mathematik; zahlreiche Auszeichnungen.



1	Einführung	2
1.1	Aufbau und Funktionsweise eines Computersystems	3
1.1.1	Die Zentraleinheit	5
1.1.2	Bildschirm, Tastatur und Maus	8
1.1.3	Weitere Ein- und Ausgabegeräte	12
1.1.4	Externe Speicher	13
1.1.5	Vernetzung	14
1.1.6	Das Betriebssystem	17
1.1.7	Fallstudie: Die Firma ProfiSoft	18
1.2	Die Informatik	18
1.3	Gliederung und Aufbau dieses Buches	25

Leser, die bereits mit Computersystemen gearbeitet haben und mit der Terminologie vertraut sind, können mit dem Kapitel 1.2 beginnen.

## 1 Einführung

Die Begriffe »Computer«, »Internet«, »Web«, »Software« und »Informatik« sind täglich in den Medien zu finden. Untersuchungen haben Folgendes ergeben /BDB 03/, /IKT-Report 03/, /TNS Emnid 03/:

- Jeder zweite Beschäftigte in Deutschland verrichtet seine Arbeit überwiegend am Computer.
- Jeder zweite Arbeitsplatz in Deutschland ist inzwischen vernetzt.
- 93 Prozent der Unternehmen in Deutschland haben einen Internet-Zugang, 75 Prozent präsentieren sich im Internet.
- Mangelnde IT-Kenntnisse von 1/3 der Mitarbeiter erweisen sich als Wachstumshindernis.
- 50,1 Prozent aller Deutschen über 14 Jahre sind 2003 online, das sind 32,1 Millionen Menschen.
- Fast 30 Millionen Konten wurden Ende 2002 in Deutschland online geführt. Damit stieg deren Zahl im Vergleich zum Vorjahr um 50 Prozent.
- Jeder zweite Arbeitslose nutzt das Internet.

Diese Zahlen deuten bereits an, dass die Informatik und insbesondere die Software-Technik heute eine große volkswirtschaftliche Bedeutung haben /BMFT 94/, /GI 87/, /Necker 94/, /Schmid, Broy 00/:

- Die Informatik zusammen mit der Informationstechnik bildet das zentrale Innovationsgebiet des 21. Jahrhunderts: »Sie wird Lebensweise, Ausbildung, Arbeit und Freizeit verändern, indem sie die Methoden des Geschäftslebens, der Forschung und der Kommunikation revolutioniert. Die Allgegenwärtigkeit der aus der Verschmelzung von Rechnern, Telefon und Medien entstehenden multimedialen Kommunikationsplattformen befreit von Ortsbindungen, ermöglicht kontinentübergreifende Kooperationen und macht neueste Erkenntnisse unmittelbar weltweit verfügbar« /Schmid, Broy 00/.
- Die Informatik hat sich in weniger als drei Jahrzehnten aus einem kleinen wissenschaftlichen Kern zu einem entscheidenden und für weite Teile unserer Wirtschaft und Gesellschaft bedeutenden Faktor entwickelt.
- Die Informatik ist zu einer Grundlagen- und Querschnittsdisziplin für die meisten Entwicklungen in Wissenschaft und Forschung, in Wirtschaft und Technik geworden.
- Software entwickelt sich zu einem eigenständigen Wirtschaftsgut. Sie ist Bestandteil der meisten hochwertigen technischen Produkte und Dienstleistungen. In einigen Bereichen, wie Banken und Versicherungen, werden nahezu alle Dienstleistungen durch den Einsatz von Software erbracht.
- Der Software-Anteil als integraler Produktbestandteil nimmt ständig zu, z.B. in der Telekommunikationsindustrie, im Auto-

mobilbau, im Maschinen- und Anlagenbau, in der Medizintechnik oder der Haushaltselektronik.

- Anlagen und Geräte werden von Software gesteuert. Sie prägt damit zunehmend sowohl die Funktionalität als auch die Qualität der Erzeugnisse.
- In exportorientierten Branchen der deutschen Wirtschaft übersteigt der Software-Anteil an der Wertschöpfung der Produkte häufig die 50-Prozent-Marke. In der digitalen Vermittlungstechnik entfallen bis zu 80 Prozent der Entwicklungskosten auf Software.
- »Information-Highways« werden einen größeren Einfluss auf die zukünftige Entwicklung von Wirtschaft und Gesellschaft haben, als es die physikalische Infrastruktur wie Schienen-, Elektrizitäts- oder Telefonnetze hatten.
- »Information« wird noch viel stärker zum entscheidenden »Produktionsfaktor« und entscheidet über die zukünftige Wettbewerbsfähigkeit der wichtigsten Industriebranchen.

Um eine intuitive Vorstellung von Computersystemen zu vermitteln und um auf klar definierten Begriffen aufzusetzen, werden im folgenden Kapitel zunächst der Aufbau, die Funktionsweise und die Terminologie von Computersystemen eingeführt und erklärt.

Anschließend wird versucht, eine Vorstellung von der Informatik als Wissenschaft zu vermitteln. Insbesondere wird auf Definitionen, Gegenstandsbereiche, Teildisziplinen und das Selbstverständnis der Informatik eingegangen.

Aufbauend auf diesen beiden Kapiteln werden dann die Gliederung und der Aufbau dieses Buches beschrieben.

## 1.1 Aufbau und Funktionsweise eines Computersystems

**Computer** sind technische Geräte, die umfangreiche Informationen mit hoher Zuverlässigkeit und großer Geschwindigkeit automatisch verarbeiten und aufbewahren können.

Computer sind *nicht* vergleichbar mit Automaten, denen man im täglichen Leben begegnet: Kaffeeautomaten, Fahrkartenautomaten, Zigarettenautomaten. Solche Automaten erfüllen nur eine festgelegte Funktion, die durch die Konstruktion bestimmt ist. Ein Zigarettenautomat kann z.B. keinen Kaffee ausgeben.

Im Gegensatz zu einem Automaten mit festgelegten Aktionen hat ein Computer einen wesentlichen Vorteil: Man kann ihm die Vorschrift, nach der er arbeiten soll, jeweils neu vorgeben. Beispielsweise kann man ihm einmal eine Arbeitsanweisung zur Berechnung der Lohnsteuer eingeben. Der Computer kann dann die Lohnsteuer berechnen. Oder man kann ihm eine Handlungsanleitung zur Steuerung von Verkehrsampeln angeben. Er ist dann in der Lage, den

Die Bezeichnung **Algorithmus** geht zurück auf den arabischen Schriftsteller Abu Dshafar Muhammed Ibn Musa *al-Khwarizmi*. Er lebte um 825 n. Chr. in der Stadt Khiva im heutigen Usbekistan, die damals Khwarizm hieß und als Teil des Namens verwendet wurde. Er beschrieb die Erbschaftsverhältnisse, die sich ergaben, wenn ein wohlhabender Araber starb, der bis zu vier Frauen in unterschiedlichem Stand und eine Vielzahl von Kindern besaß. Dazu verwendete er algebraische Methoden und schrieb ein Lehrbuch mit dem Titel »Kitab al jabr w'almuqabalah« (Regeln zur Wiederherstellung und zur Reduktion), wobei die Übertragung von Gliedern einer Gleichung von einer zur anderen Seite des Gleichheitszeichens gemeint ist. Der Begriff Algebra leitete sich aus dem Titel des Lehrbuchs ab. Aus dem Namen des Schriftstellers wurde *algorism* und daraus *Algorithmus*.

## LE 1 1.1 Aufbau und Funktionsweise eines Computersystems

Verkehrsfluss einer Stadt zu regeln. Solche Vorschriften, Arbeitsanweisungen oder Handlungsanleitungen bezeichnet man als **Algorithmen** (Singular: **Algorithmus**).

Damit ein Computer Algorithmen ausführen kann, müssen sie in einem eindeutigen und präzisen Formalismus beschrieben werden. Ist dies der Fall, dann bezeichnet man solche Algorithmen als **Programme**. Da ein Computer ohne ein Programm nicht arbeiten kann, spricht man von **Computersystemen**, wenn das technische Gerät Computer *und* die Programme zur Steuerung des Computers gemeint sind.

Alle materiellen Teile eines Computersystems bezeichnet man auch als **Hardware** (harte Ware), alle immateriellen Teile, z.B. Programme, eines Computersystems als **Software** (weiche Ware). Analogien zu diesen Begriffen sind beispielsweise:

Hardware – Musikinstrument – Schienennetz,

Software – Komposition – Zugfahrplan.

Alles, was ein Computersystem kann, kann ein Mensch im Prinzip auch. Ein Computersystem hat gegenüber dem Menschen jedoch drei wesentliche Vorteile:

- Hohe Speicherkapazität

In einem Computersystem können Unmengen von Informationen – Millionen bis Billionen Zeichen – aufbewahrt werden. Gesuchte Informationen können sehr schnell wiedergefunden werden.

- Hohe Geschwindigkeit

Ein Computersystem addiert heute in einer Sekunde rund zehn Millionen Zahlen. Nimmt man an, ein Mensch würde in einer Sekunde zwei solcher Zahlen addieren, dann hätte er nach einem Jahr ununterbrochener Tätigkeit ca. 32 Millionen Zahlen zusammengezählt. Ein Computersystem schafft das in drei Sekunden. Der schnellste heute verkaufte Computer schafft sogar 35 Billionen Rechenoperationen pro Sekunde ( $10^{12}$  Operationen pro Sekunde = 1 Tera-Operation pro Sekunde).

- Hohe Zuverlässigkeit

Ein weiterer Vorzug eines Computersystems ist seine hohe Zuverlässigkeit. Während der Mensch, insbesondere bei monotonen Aufgaben, schnell ermüdet und Fehler macht, kennt ein Computersystem solche Schwächen nicht.

Der Begriff Software ist noch umfassender als der Begriff Programm.

Software **Software** (SW) sind Programme, zugehörige Daten und notwendige Dokumentation, die es zusammengefasst erlauben, mithilfe eines Computers Aufgaben zu erledigen. Synonym zu Software werden oft auch die Begriffe **Software-System** und **Software-Produkt** verwendet.

Software gliedert man oft in Anwendungssoftware und Systemsoftware.

Die 500 schnellsten Computer:  
[www.top500.org](http://www.top500.org)



### 1.1.1 Die Zentraleinheit LE 1

**Systemsoftware**, auch Basissoftware genannt, ist Software, die für eine spezielle Hardware oder eine Hardwarefamilie entwickelt wurde, um den Betrieb und die Wartung dieser Hardware zu ermöglichen bzw. zu erleichtern. Zur Systemsoftware zählt man immer das Betriebssystem, in der Regel aber auch Compiler, Datenbanken, Kommunikationsprogramme und spezielle Dienstprogramme (siehe unten). Systemsoftware orientiert sich grundsätzlich an den Eigenschaften der Hardware, für die sie geschaffen wurde und ergänzt normalerweise die funktionalen Fähigkeiten der Hardware.

Systemsoftware

**Anwendungssoftware**, auch Applikationssoftware (*application software*) genannt, ist Software, die Aufgaben des Anwenders mithilfe eines Computersystems löst. Beispiele dafür sind Textverarbeitungs-Software, Tabellenkalkulation, Zeichenprogramme. Anwendungssoftware setzt in der Regel auf der Systemsoftware der verwendeten Hardware auf bzw. benutzt sie zur Erfüllung der eigenen Aufgaben.

Anwendungssoftware

Anwendungssoftware, Systemsoftware und Hardware bilden zusammen ein **Computersystem**.

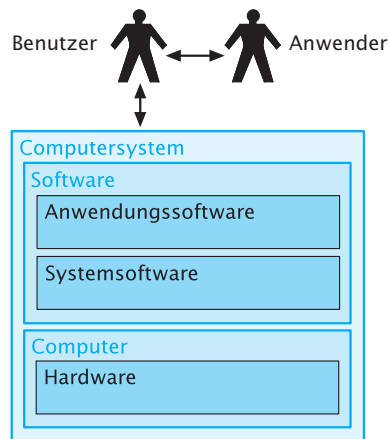
Computersystem

Als **Anwender** werden alle Angehörigen einer Institution oder organisatorischen Einheit bezeichnet, die ein Computersystem zur Erfüllung ihrer fachlichen Aufgaben einsetzen. Sie benutzen die Ergebnisse der Anwendungssoftware oder liefern Daten, die die Anwendungssoftware benötigt.

Anwender

**Benutzer** sind nur diejenigen Personen, die ein Computersystem unmittelbar einsetzen und *bedienen*, oft auch Endbenutzer oder Endanwender genannt.

Abb. 1.1-1 zeigt nochmals grafisch die Zusammenhänge.



Benutzer

Abb. 1.1-1:  
Begriffe und ihre  
Zusammenhänge

### 1.1.1 Die Zentraleinheit

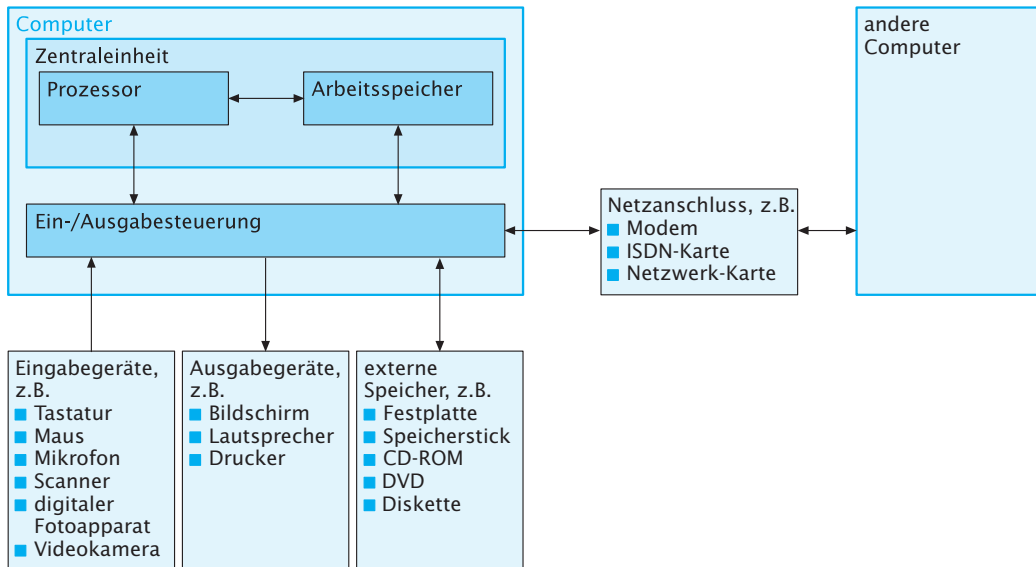
Ein Computer besteht aus

- einer **Zentraleinheit** (*central unit*) und
- einer Ein-/Ausgabesteuerung.

In der Zentraleinheit werden die Programme abgearbeitet. Sie besteht aus

- dem Prozessor – auch CPU (*central processing unit*) genannt – und
- dem Arbeitsspeicher (Abb. 1.1-2).

## LE 1 1.1 Aufbau und Funktionsweise eines Computersystems



Peripheriegeräte

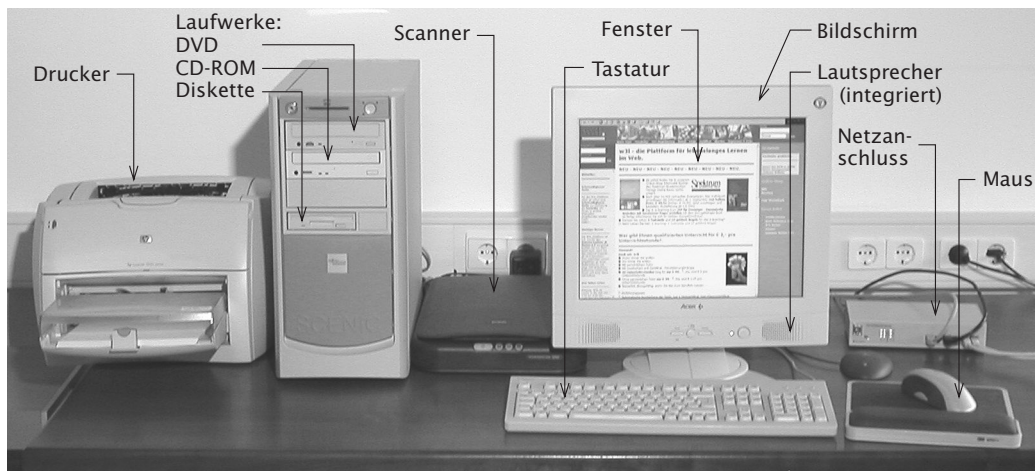


Abb. 1.1-2:  
Aufbau und  
Komponenten eines  
Computers

Die Zentraleinheit kommuniziert über eine Ein-/Ausgabesteuerung mit den Peripheriegeräten (Ein-/Ausgabegeräte, externe Speicher, Netzanschluss).

Arbeitsspeicher

Der **Arbeitsspeicher**, auch **RAM** (*random access memory* = Direktzugriffsspeicher) genannt, dient dazu, Informationen kurzfristig aufzubewahren und bei Bedarf wieder zur Verfügung zu stellen. Die Aufbewahrung von Informationen erfolgt nur *kurzfristig*, da die Informationen nach dem Ausschalten des Computers, d.h. nach dem Abschalten der Stromzufuhr, wieder verloren gehen. Dies hängt mit dem technischen Aufbau von Arbeitsspeichern zusammen.

Für die *langfristige* Aufbewahrung von Informationen werden deshalb **externe Speicher** (siehe Abschnitt 1.1.2) verwendet, die die Informationen ähnlich wie bei einem Tonband oder einer CD aufbewahren.

Der Hauptvorteil des Arbeitsspeichers liegt darin, dass Informationen wesentlich schneller hineingegeben und wieder herausgeholt werden können als bei externen Speichern.

Die kleinste Einheit eines Arbeitsspeichers bezeichnet man als **Speicherzelle**. Eine Speicherzelle kann im Allgemeinen eine Zahl oder mehrere Zeichen aufbewahren. Jede Speicherzelle besitzt einen Namen (Adresse). Durch die Angabe des Namens wird genau die dem Namen zugeordnete Speicherzelle angesprochen. Der Arbeitsspeicher eines Computers ist *nicht unendlich* groß, d.h., er verfügt *nicht* über unendlich viele Speicherzellen.

Speicherzelle

Die Informationsmenge, die in einem **Speicher** aufbewahrt werden kann, wird als dessen **Speicherkapazität** bezeichnet. Die Maßeinheit für Kapazität heißt **Byte**. In einem Byte kann bei Texten in der Regel ein Zeichen gespeichert werden. Ein Byte wird zu folgenden Einheiten zusammengefasst:

Speicherkapazität  
Byte

1 KB = 1 Kilo-Byte = 1.024 Bytes

1 MB = 1 Mega-Byte = 1.024 KB = 1.048.576 Bytes  $\approx$  1 Million Bytes

1 GB = 1 Giga-Byte = 1.024 MB  $\approx$  1 Milliarde Bytes.

Im Unterschied zum normalen Kilobegriff, der 1.000 bedeutet, hat Kilo hier den Faktor 1.024. Die Ursache dafür ist, dass Bytes im Binär- bzw. Dualsystem dargestellt werden.  $2^{10}$  ergibt 1.024 und nicht 1.000.

Ein Byte besteht wiederum aus acht Binärzeichen, abgekürzt **Bit** (*binary digit*) genannt. Ein Bit kann einen zweielementigen Zeichenvorrat darstellen, z.B. Markierung – keine Markierung, wahr – falsch, eins – null.

In Abhängigkeit von der Größe eines Computers hat ein Arbeitsspeicher heute in der Regel eine Kapazität zwischen acht MB und mehreren hundert MB.

Externe Speicher haben folgende Speicherkapazitäten:

- Diskette: 1,44 MB
- Speicherstick: 64–512 MB
- CD-ROM: 650 MB
- DVD: 4,7 GB bis 40 GB
- Festplatte: 2 GB bis ca. 10 GB (an einen Computer können mehrere Festplatten angeschlossen werden, sodass Speicherkapazitäten bis 200 GB und mehr erreicht werden).

Um einen Eindruck von der Größenordnung der Speicherkapazität zu bekommen, sei ein Rechenbeispiel angegeben:

Beispiel

Eine maschinengeschriebene DIN-A4-Text-Seite fasst rund 3.000 Zeichen einschließlich der Leerzeilen und Leerzeichen; benötigte Speicherkapazität: 3 KB. Für ein Buch mit 1.000 Seiten wird also eine Speicherkapazität von 3 MB benötigt.



## LE 1 1.1 Aufbau und Funktionsweise eines Computersystems

Heutige Großplattenspeicher haben eine Kapazität von bis zu 200 GB oder mehr. Auf solch einem Speicher kann daher eine ganze Bibliothek, bestehend aus 70.000 Büchern, abgespeichert werden.

Prozessor	<p>Die zweite Komponente der Zentraleinheit ist der Prozessor. Der <b>Prozessor</b> eines Computers ist in der Lage, Programme abzuarbeiten. Bei Programmen handelt es sich um Algorithmen, die stark formalisiert sind. Algorithmen dagegen können umgangssprachlich formuliert sein. Führt ein Mensch einen Algorithmus aus, dann weiß er durch seine Erfahrung auch bei nicht detaillierten Formulierungen meist, wie der Algorithmus zu verstehen ist.</p> <p>Bei dem Prozessor eines Computers handelt es sich dagegen um ein technisches Bauteil, dem in einem präzisen Formalismus mitgeteilt werden muss, was es tun soll. Den Formalismus, den ein Prozessor versteht, bezeichnet man als Programmiersprache, genauer gesagt als Maschinensprache.</p>
Programmiersprache	<p>Im Gegensatz zur Umgangssprache ist eine <b>Programmiersprache</b> eine formalisierte Sprache,</p> <ul style="list-style-type: none"><li>■ deren Sätze aus einer Aneinanderreihung von Zeichen eines festgelegten Zeichenvorrates entstehen,</li><li>■ deren Sätze aufgrund einer endlichen Menge von Regeln gebildet werden müssen (Syntax) und</li><li>■ die die Bedeutung jedes Satzes festlegt (Semantik).</li></ul>
Programm	<p>Ein <b>Programm</b> ist deshalb ein Algorithmus, formuliert in einer Programmiersprache.</p> <p>Der Prozessor verfügt zum Verarbeiten von Informationen über folgende Fähigkeiten:</p> <ul style="list-style-type: none"><li>■ Informationen aus dem Arbeitsspeicher in den Prozessor transportieren,</li><li>■ Informationen vergleichen, addieren, subtrahieren, multiplizieren, dividieren, verknüpfen,</li><li>■ Informationen aus dem Prozessor in den Arbeitsspeicher transportieren.</li></ul> <p>Die prinzipielle Arbeitsweise der Zentraleinheit zeigt Abb. 1.1-3.</p>

### 1.1.2 Bildschirm, Tastatur und Maus

**Bildschirm, Tastatur** und **Maus** sind die wichtigsten Geräte, über die der Computerbenutzer mit dem Computersystem kommuniziert (Abb. 1.1-2).

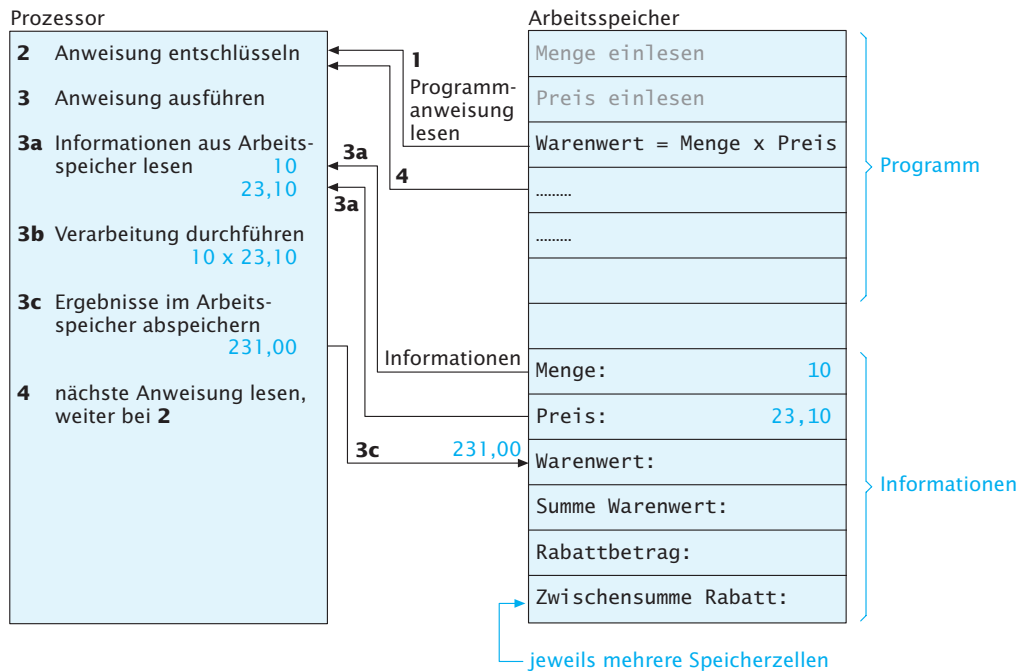
Bildschirm	<p>Auf einem Bildschirm können Informationen angezeigt werden. Heute üblich sind gerasterte Grafikbildschirme. Sie ermöglichen die Darstellung von Zeichen, Grafiken, Bildern und Filmen.</p>
Pixel	<p>Die kleinsteanzeigeeinheit ist hierbei ein bildpunkt, meist <b>Pixel</b> (<i>picture element</i>) genannt. Wie fein, d.h. mit welcher <b>Auflösung</b> grafische Darstellungen angezeigt werden können, hängt von der</p>



## 1.1.2 Bildschirm, Tastatur und Maus LE 1

Ein Programm besteht aus eine Reihe von Anweisungen. Im folgenden Beispiel wird angenommen, dass für eine Rechnungserstellung der Warenwert pro Artikelposition aus Menge und Preis zu berechnen ist (Fakturierung). Es wird angenommen, dass der Benutzer über die Tastatur bereits die Menge und den Preis eingegeben hat und dass diese Werte sich bereits im Arbeitsspeicher befinden. Sowohl das Programm zur Fakturierung als auch die Eingabe- und Ausgabeinformationen, die bei der Ausführung des Programms benötigt bzw. erzeugt werden, werden im Arbeitsspeicher aufbewahrt.

Der Prozessor steuert die Informationsverarbeitung und führt arithmetische, logische und Transportoperationen aus. Transportoperationen sind die Operationen, die den Transport der Informationen aus dem Arbeitsspeicher in den Prozessor und umgekehrt bewirken.



Der Prozessor liest eine Programmanweisung nach der anderen und führt jede Programmanweisung mit den im Arbeitsspeicher aufbewahrten Informationen aus. Zwischenergebnisse und Ergebnisse werden vom Prozessor im Arbeitsspeicher abgelegt.

Die Idee, Programme und Informationen in einem gemeinsamen Speicher aufzubewahren, stammt von Johann von Neumann (siehe Kurzbiografie am Kapitelanfang). Computer mit einem Prozessor bezeichnet man ihm zu Ehren als von-Neumann-Computer.

Anzahl der Pixel ab. Gute Grafikeigenschaften erhält man ab 800 mal 600 Bildpunkten (SVGA-Auflösung).

Eine Computertastatur besteht aus Tasten für die Eingabe von Informationen (Buchstaben, Ziffern, Sonderzeichen) sowie einigen Sondertasten zur Computerbedienung.

Die wichtigsten sind:

Auslöse- oder Eingabe-Taste (*carriage return, enter*): ↵

Sie dient dazu, eine Informationseingabe abzuschließen.

Steuerungs-Taste (*control*): CTRL oder STRG

**Abb. 1.1-3:**  
**Prinzipielle**  
**Arbeitsweise der**  
**Zentraleinheit**

Tastatur

## LE 1 1.1 Aufbau und Funktionsweise eines Computersystems

Diese Taste, zusammen mit anderen Tasten gleichzeitig gedrückt, bewirkt bestimmte Aktionen. Beispiel: CTRL- und p-Taste bewirkt bei manchen Systemen den Ausdruck von Informationen.

Am oberen Rand einer Computertastatur befinden sich oft so genannte **frei programmierbare Funktionstasten**, meist bezeichnet mit F1 bis F12. Derjenige, der ein Programm schreibt, kann jede dieser Tasten mit einer bestimmten Bedeutung belegen. Wird das Programm vom Computer ausgeführt und die Taste wird gedrückt, dann wird eine entsprechende Funktion ausgelöst. Mit der Funktionstaste F1 wird im Allgemeinen eine Hilfefunktion aufgerufen.

Zusätzlich enthält jede Computertastatur so genannte **Cursor-Tasten**. Ein *Cursor* (Läufer), auch Schreibmarke genannt, ist eine oft blinkende Markierung (meist ein senkrechter Strich) auf dem Bildschirm, der die momentane Bearbeitungsposition anzeigt. Gibt man ein Zeichen mit der Tastatur ein, dann wird es im Allgemeinen an der Stelle eingefügt, an der der *Cursor* steht. Mithilfe der *Cursor*-Tasten kann der *Cursor* auf dem Bildschirm bewegt werden.

Da Grafikbildschirme aus mehreren hunderttausend Bildpunkten bestehen, ist ein schnelles und genaues Positionieren mit *Cursor*-Tasten nicht möglich.

Maus Als Alternativen wurden so genannte **Zeiginstrumente** entwickelt, die es ermöglichen, schnell direkt oder indirekt auf Bildpunkte oder Bildbereiche zu zeigen. Am häufigsten wird hierfür heute die so genannte **Maus** verwendet. Eine Maus ist ein kleines handliches Kästchen mit einer, zwei oder drei Drucktasten auf der Oberfläche und einer Rollkugel auf der Unterseite (Abb. 1.1-2).

Die Maus wird auf eine glatte Oberfläche gelegt und kann dann leicht hin- und herbewegt werden. Die Bewegung der Maus wird dabei von einem Maus-Zeiger, meist durch einen Pfeil dargestellt, auf dem Bildschirm nachvollzogen. Durch Betätigen einer Taste (Anklicken) kann man bestimmte Aktionen auslösen.

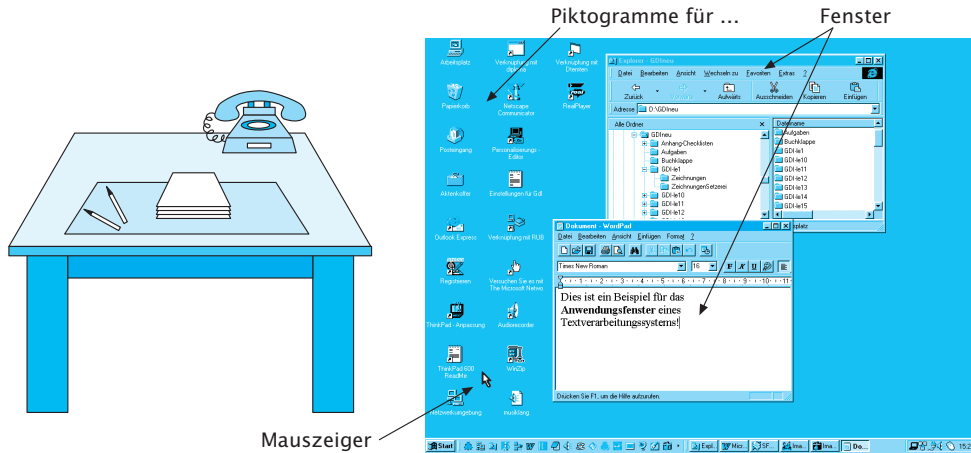
Durch die Verwendung eines Grafikbildschirms ist es auch möglich, die Informationsdarstellung auf dem Bildschirm so zu gestalten, dass die Bedienung des Computers besonders einfach wird.

GUI = graphical user interface Der Benutzer sieht eine **grafische Benutzungsoberfläche (GUI)**, die einer physikalischen Arbeitsoberfläche eines Schreibtisches nachgebildet ist. Man spricht auch von einer »elektronischen **Arbeitsoberfläche, desktop**« genannt.

Arbeitsoberfläche (desktop) Auf dieser Arbeitsoberfläche können Objekte (Aufträge, Bestellungen, Rechnungen, Memos usw.), Hilfsmittel (Bleistift, Radiergummi, Büroklammern usw.), Geräte (Telefon, Posteingangs- und -ausgangskorb) und Anwendungen (Adressverwaltung, Textverarbeitungssystem usw.), die die Bearbeitung der Objekte ermöglichen, in Form von **Piktogrammen** bzw. Ikonen (**icons**) dargestellt werden.

Piktogramme (icons) Ein Piktogramm ist eine grafisch abstrakte Darstellung von Objekten, Geräten, Funktionen, Anwendungen oder Prozessen auf dem

## 1.1.2 Bildschirm, Tastatur und Maus LE 1



Bildschirm. Abb. 1.1-4 zeigt die Analogie zwischen einem »klassischen« Schreibtisch-Arbeitsplatz und der Arbeitsoberfläche eines Computer-Arbeitsplatzes.

Der Benutzer kann auf der elektronischen Arbeitsoberfläche eine durch ein Piktogramm repräsentierte Anwendung aktivieren. In der Regel geschieht dies durch einen Doppelklick mit der Maus auf das entsprechende Piktogramm. Die Anwendung öffnet dann ein Anwendungsfenster auf der Arbeitsoberfläche.

Ein **Fenster** entspricht einem Ausschnitt aus der Arbeitsumgebung des Benutzers und ermöglicht die Bearbeitung entsprechender Objekte. Beispielsweise können nach dem Start eines Textsystems Textobjekte (Briefe, Memos, Dokumente) bearbeitet werden.

Der Benutzer kann mehrere Anwendungen aktivieren, sodass mehrere Anwendungsfenster auf der Arbeitsoberfläche sichtbar sind. Anwendungsfenster stellen die Benutzungsoberfläche der jeweiligen Anwendung dar. Ein Anwendungsfenster kann wiederum aus mehreren Fenstern bestehen.

Die Verwaltung der Arbeitsoberfläche, die Bereitstellung einer anwendungsorientierten Grundfunktionalität sowie die Koordination mit den verschiedenen Anwendungen übernimmt ein **GUI-System**. Das verwendete GUI-System beeinflusst ganz wesentlich die Gestaltung der Anwendungsoberflächen.

Die auf der Arbeitsoberfläche verfügbaren Elemente »Fenster« und »Piktogramme« müssen vom Benutzer manipuliert werden können, damit er seine Aufgaben erledigen kann. Wenn auf der Arbeitsoberfläche bereits versucht wird, eine physikalische Arbeitsumgebung nachzubilden, dann liegt es natürlich nahe, auch die Arbeitsweise auf der elektronischen Arbeitsoberfläche der physikalischen Arbeitsweise in übertragener Form anzupassen.

*Abb. 1.1-4:  
Vergleich einer  
traditionellen  
Arbeitsoberfläche  
mit einer elektroni-  
schen Arbeitsober-  
fläche*

GUI-System

## LE 1 1.1 Aufbau und Funktionsweise eines Computersystems

**Beispiel** Ein Mitarbeiter, der ein Stück Papier nicht mehr benötigt, ergreift es mit der Hand und wirft es in den Papierkorb.

Selektieren, Ziehen, Loslassen  
(*pick, drag & drop*) Auf der elektronischen Arbeitsoberfläche kann der Benutzer analog das Stück Papier, repräsentiert durch ein Piktogramm, mit der Maus selektieren (*pick*), mit gedrückter Maustaste das Papierpiktogramm über die Arbeitsoberfläche auf das Papierkorbpiktogramm bewegen (*drag*) und dann die Maustaste loslassen (*drop*). Das Papierpiktogramm verschwindet. Das Papierkorbpiktogramm ändert seine Form, um anzuzeigen, dass der Papierkorb etwas enthält (Abb. 1.1-5).



Abb. 1.1-5:  
Selektieren, Ziehen  
und Loslassen  
(Beispiel)

Benötigt der Mitarbeiter das Stück Papier aus dem Papierkorb doch noch einmal, dann kann er es aus dem Papierkorb zurückholen. Dies muss auch beim elektronischen Papierkorb möglich sein, sonst stimmt die Analogie nicht. Ist dies nicht möglich, dann muss ein Reißwolf anstelle des Papierkorbs durch ein Piktogramm dargestellt werden.

Bei der **direkten Manipulation** werden vom Benutzer – in gewisser Analogie zur Arbeitsweise in einer physikalischen Arbeitsumgebung (z.B. Büro) – Arbeitsobjekte (z. B. Dokumente) unmittelbar visuell identifiziert, selektiert (»zur Hand genommen«) und bearbeitet. Die Bedienungstechnik »Selektieren, Ziehen und Loslassen« (*pick, drag & drop*) ist ein Beispiel für eine direkte Manipulation.

direkte  
Manipulation

### 1.1.3 Weitere Ein- und Ausgabegeräte

Neben Texten werden in vielen Anwendungen auch Grafiken, Fotos und Bilder verwendet. Um diese Informationen im Computer speichern zu können, werden sie durch einen **Scanner** (Abtaster) oder einen digitalen Fotoapparat aufgenommen, digitalisiert und als Bitmuster dem Computersystem zur Verfügung gestellt.

Scanner

Videofilme können ebenfalls digitalisiert, gespeichert und auf dem Bildschirm abgespielt werden.

digitale  
Videokamera

Will man Ausgabeinformationen vom Computer nicht nur auf dem Bildschirm lesen, sondern »schwarz auf weiß« bzw. in Farbe besitzen, dann benötigt man als Ausgabegerät einen **Drucker**. Es gibt heute eine Vielzahl von Druckern, die ein breites Leistungs- und Preisspektrum abdecken. Die Auswahl eines geeigneten Druckers hängt vom Einsatzgebiet ab. Aus dem Einsatzgebiet lassen sich Anforderungen an den Drucker ableiten. Häufig verwendet werden heute

Drucker

- Laserdrucker und
- Tintenstrahldrucker.

Mikrofon und Lautsprecher ermöglichen den Einsatz gesprochener Sprache in Computersystemen.

Mikrofon,  
Lautsprecher  
gesprochene  
Sprache

In verschiedenen Formen wird die gesprochene Sprache zukünftig in der Kommunikation mit einem Computersystem und über ein Computersystem verstärkt möglich sein. Die Bearbeitung gespro-

chener Sprache lässt sich unterteilen in Sprachspeicherung, Spracherkennung und Sprachausgabe.

Bei der **Sprachspeicherung** wird die Sprache über ein Mikrofon aufgenommen. Die Sprache wird digitalisiert, die Daten werden komprimiert und dann auf einem Speichermedium aufbewahrt.

Sprachspeicherung

Die Sprachspeicherung kann im Büro für folgende Anwendungen eingesetzt werden:

- Verwendung als **Diktiergerät**,
- Versenden von **Sprachmitteilungen** (*voice mailing*),
- **Sprachanmerkungen** in Dokumenten.

Wie bei einem Diktiergerät kann Sprache aufgezeichnet und über einen Lautsprecher abgehört werden. Zusätzlich können Einfügungen in bereits aufgezeichnete Sprachinformationen vorgenommen werden. Ähnlich wie Dokumente können auch Sprachinformationen in Form von **Sprachmitteilungen** über elektronische Post an einen oder mehrere Empfänger versendet werden.

**Spracherkennung** bedeutet, dass gesprochene Informationen vom Computersystem erkannt werden. Spricht der Benutzer z.B. das Wort »Posteingang«, dann muss dieses Wort vom System erkannt und anschließend beispielsweise die eingegangene Post angezeigt werden. Die Spracherkennung eröffnet neue Anwendungen, insbesondere die direkte Sprach-Text-Umsetzung (»automatische Schreibmaschine«, »Sprechschreiber«).

Spracherkennung

Die vollsynthetische **Sprachausgabe** erlaubt es, einen beliebigen im Computersystem befindlichen Text vorlesen zu lassen. In mehreren Schritten wird aus einem vorliegenden Text eine Sprachausgabe aufgebaut und synthetisiert, ohne dass vorher eine menschliche Stimme als Vorlage oder Referenz vorhanden war.

Sprachausgabe

Zunehmend an Bedeutung gewinnen **Multimedia**-Anwendungen, die Texte, Daten, Grafiken, Bilder, Filme, Töne und gesprochene Sprache kombinieren.

Multimedia

### 1.1.4 Externe Speicher

Arbeitsspeicher besitzen sowohl aus technischen als auch aus Kostengründen eine Kapazität in der Größenordnung von vier MB bis zu mehreren hundert MB. Es ist daher nicht möglich, umfangreiche Informationsbestände und eine Vielzahl von Programmen im Arbeitsspeicher über längere Zeit hinweg aufzubewahren. Außerdem gehen alle Informationen im Arbeitsspeicher verloren, wenn der Computer ausgeschaltet wird.

Im Arbeitsspeicher befinden sich daher nur die Programme und Informationen, mit denen gerade gearbeitet wird. Für die langfristige Aufbewahrung von Programmen und Informationen hat man besondere Speicher – die **externen Speicher** – entwickelt, die wesentlich preiswerter als Arbeitsspeicher sind und eine höhere Speicher-

externe Speicher

## LE 1 1.1 Aufbau und Funktionsweise eines Computersystems

kapazität besitzen. Der Begriff *externe Speicher* deutet an, dass diese Speicher nicht zur Zentraleinheit gehören.

Es gibt verschiedene externe Speichermedien. Sie unterscheiden sich in der Zugriffsgeschwindigkeit, d.h. wie schnell Informationen wiedergefunden werden, in der Speicherkapazität und im Preis. Sie lassen sich in fünf große Klassen einteilen:

- **Diskettenspeicher** (geringe Kapazität, langsam),
- **Speicherstick** (mittlere Kapazität, schnell),
- **Plattenspeicher** (hohe Kapazität, schnell),
- **CD-ROM-Speicher** (mittlere Kapazität, mittlere Zugriffsgeschwindigkeit),
- **DVD-Speicher** (hohe Kapazität, mittlere Zugriffsgeschwindigkeit).

Die Speichermedien Diskette, Platte, CD-ROM und DVD befinden sich in Laufwerken, die das entsprechende Medium an einem Schreib-/Lesekopf zum Schreiben und Lesen der Informationen vorbeibewegen.

Dateien Informationen werden auf externen Speichern in **Dateien (files)** abgelegt. Dateien sind vergleichbar mit Karteikästen bei der Ablage manueller Informationen. Um eine Datei speichern zu können, muss sie einen Namen erhalten. Dieser Name wird in einem **Inhaltsverzeichnis (directory)** auf dem Speichermedium vermerkt. Ebenso die Größe der Datei in Bytes. In diesem Inhaltsverzeichnis sind die Namen aller gespeicherten Dateien enthalten.

Ordner, Verzeichnisse Dateien können zu Ordnern bzw. Verzeichnissen zusammengefasst werden, Ordner können wiederum Ordnern zugeordnet werden. Dadurch ist es möglich, Ordnerhierarchien aufzubauen, z.B. Informatik/Grundlagen/LE1.doc, wobei LE1.doc der Dateiname ist, Grundlagen der zugehörige Ordner. Der Ordner Grundlagen gehört wiederum zum Ordner Informatik.

Hinweis: Der Aufbau von Dateinamen hängt vom verwendeten Betriebssystem (siehe Abschnitt 1.1.6) ab. Die hier beschriebene Zweiteilung ist in dem Betriebssystem MS-DOS, das heute Bestandteil von Windows ist, vorgeschrieben. In anderen Betriebssystemen ist eine Unterteilung in ein oder mehrere Teile möglich, aber nicht vorgeschrieben.

Um eine Datei eindeutig von allen anderen Dateien in einem Computersystem unterscheiden zu können, muss die gesamte Hierarchie, **Pfad** genannt, angegeben werden, wie oben im Beispiel.

Oft gliedert man Dateinamen in zwei Teile. Bei dem ersten Teil handelt es sich um den eigentlichen, frei wählbaren Namen wie LE1. Der zweite Teil ist die Dateinamen-Erweiterung, Suffix oder *extension* genannt, die den Typ der Datei angibt, z.B. txt für eine Textdatei, doc für ein *Word*-Dokument, wav für eine Audio-Datei, wmf für eine Grafik-Datei, bmp für eine Bild-Datei. Der Dateinamen-Suffix wird auch dazu verwendet, um zu einer Datei automatisch die zugehörige Anwendungssoftware zu starten.

### 1.1.5 Vernetzung

Computersysteme sind heute in der Regel vernetzt, d.h., sie können mit anderen Computersystemen Informationen austauschen. Damit Computersysteme elektronisch kommunizieren können, müssen sie

im Allgemeinen physikalisch durch Kabel oder über Funk miteinander verbunden sein. Die Kabel und die Funkstrecken, die die Verbindung herstellen, werden in ihrer Gesamtheit als **Netz** bezeichnet. Zu dem Kabel als Hardware bzw. den Sende- und Empfangsstationen bei Funk kommt noch die Netz-Software hinzu, die mit dafür sorgt, dass die elektronische Informationsübermittlung stattfinden kann. Zwei wichtige Netze sind

- das Intranet und
- das Internet.

Ein **Intranet** verbindet Computersysteme eines Unternehmens oder einer Organisation.

Das **Internet** verknüpft weltweit Intranets oder einzelne Computersysteme miteinander. Auf das Internet wird ausführlich im Kapitel 1.4 eingegangen.

Kapitel 1.4

Es gibt mehrere technische Möglichkeiten, um Netzwerke zu realisieren. In Abhängigkeit von der notwendigen Übertragungsgeschwindigkeit der Informationen können folgende Übertragungsmedien benutzt werden:

- Telefon- oder ISDN-Leitungen für geringe Übertragungsgeschwindigkeiten zwischen 14.400 Bit/s und 64.000 Bit/s. Das entspricht 1.800 Zeichen/s bzw. 8.000 Zeichen/s.
- ADSL-Zugang über Telefon- oder ISDN-Kupfer-Leitungen für Geschwindigkeiten zwischen 768 KBit/s und 3072 KBit/s (zum Herunterladen) sowie 128 KBit/s und 384 KBit/s (zum Heraufladen).
- Verdrehte Zweidrahtleitungen (*twisted pair*) oder Koaxialkabel für mittlere Übertragungsgeschwindigkeiten. Die Geschwindigkeiten liegen zwischen 10 MBit/s, das entspricht etwa 1,25 Millionen Zeichen/s, und 100 MBit/s.
- Lichtwellenleiter bzw. Glasfasern für hohe Übertragungsgeschwindigkeiten (100 MBit/s bis hin zu 500 MBit/s). Solche Lichtwellenleiter werden zunehmend für die Verkabelung von Telekommunikations-Netzen und lokalen Netzen verwendet.

ISDN steht für »integriertes digitales Fernmelddenetz« (*integrated services digital network*).

ADSL steht für »Asymmetric Digital Subscriber Line«

Die Übertragung von Informationen in Netzen erfolgt **paketweise**, d.h., die zu übertragenden Informationen werden in ein oder mehrere Pakete (Abb. 1.1-6) fester oder variabler Länge aufgeteilt. Die Pakete werden dann über das Netz transportiert. Die Größe eines Paketes wird üblicherweise beschränkt, sodass die Übertragungsdauer eine bestimmte Zeit nicht überschreitet (meist im Bereich von tausendstel Sekunden). Bei Übertragungsgeschwindigkeiten im MBit/s-Bereich ergeben sich Paketgrößen im KBit-Bereich. Beispielsweise können

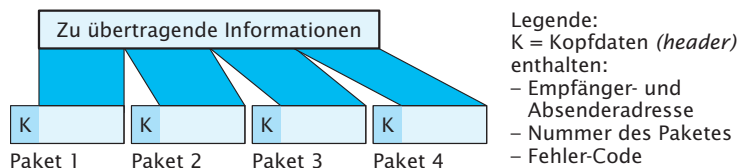


Abb. 1.1-6:  
Zum Prinzip der Informationsübertragung mit Paketen



## LE 1 1.1 Aufbau und Funktionsweise eines Computersystems

250 Zeichen = 250 Byte = 2.000 Bit = 2 KBit als ein Paket übertragen werden.

In Abhängigkeit davon, ob zu einem Zeitpunkt nur ein Paket oder mehrere Pakete parallel im Netz transportiert werden, unterscheidet man **Basisbandnetze** und **Breitbandnetze**.

In **Basisbandnetzen** werden mehrere Pakete nur zeitlich hintereinander transportiert (Zeitmultiplex).

**Breitbandnetze** können zeitlich parallel mehrere Pakete übertragen. Dies geschieht dadurch, dass das Übertragungsmedium, meist Lichtwellenleiter, mit mehreren nicht überlappenden Frequenzbändern beschickt wird (vergleichbar mit der gleichzeitigen Übertragung verschiedener Sender im Rundfunk). Auf jedem Frequenzband kann unabhängig von den anderen Frequenzbändern ein Paket übertragen werden (Frequenzmultiplex).

Durch so genannte **Übertragungsprotokolle** wird geregelt, auf welche Weise ein Computersystem Zugang zum Übertragungsmedium erhält. In Abhängigkeit vom Übertragungsmedium und dem verwendeten Übertragungsprotokoll sind für ein Computersystem unterschiedliche Netzanschluss-Systeme nötig:

Netzanschluss-System

- Ein **Modem** (Modulator/Demodulator) wird benötigt, wenn das Computersystem an ein analoges Telefonnetz angeschlossen werden soll. Es wandelt digitale Signale, d.h. Bits, in akustische um. Diese Umwandlung heißt Modulation. Beim Empfänger werden akustische Signale in digitale umgesetzt (Demodulation).
- Ein **DSL-Modem** wird benötigt, um einen ADSL-Zugang zu ermöglichen.
- Eine **ISDN-Karte** wird benötigt, wenn das Computersystem an ein digitales Telefonnetz angeschlossen werden soll. An einen ISDN-Basisanschluss können bis zu acht unterschiedliche Endgeräte angeschlossen werden, von denen wiederum zwei Geräte gleichzeitig benutzt werden können. Beim Schmalband-ISDN sind Übertragungsgeschwindigkeiten bis zu 64 KBit/s erlaubt. Langfristig wird ein Breitband-ISDN angestrebt, das ein Vielfaches dieser Übertragungsgeschwindigkeit ermöglicht.
- Eine **Netzwerk-Karte** wird benötigt, wenn das Computersystem an ein lokales Netzwerk (LAN, *local area network*) angeschlossen werden soll. Lokale Netzwerke verbinden Computersysteme über kurze Entfernungen (von einigen Metern bis maximal wenigen Kilometern), meist innerhalb einer Institution oder eines Unternehmens. Als Übertragungsmedium werden keine »öffentlichen« Leitungen verwendet.

Ein Netzwerk kann nicht nur dazu dienen, Informationen zwischen den angeschlossenen Computern auszutauschen (Informationsverbund), sondern auch dazu, Ressourcen gemeinsam zu nutzen. Ressourcen können beispielsweise Drucker und externe Speicher sein.

Hinweis: Der hier benutzte Begriff Karte bezeichnet eine technische Komponente eines Computers, die in das Gehäuse eines Computers nachträglich eingebaut werden kann, daher auch Einsteckkarte oder Einschubkarte genannt. Die Bezeichnung Karte beruht auf der Form der technischen Komponente. Sie hat eine rechteckige Form und besitzt ungefähr die Größe einer DIN-A6-Seite.

Da diese Peripheriegeräte relativ teuer sind, können sie als so genannte Server an das Netz angeschlossen werden.

Ein **Server** ist ein Gerät, das festgelegte Dienstleistungen auf Anforderung für die Teilnehmer im Netz erbringt. Die im Netz befindlichen Computer können dann Druckaufträge an den Druck-Server (es handelt sich dabei z.B. um einen schnellen Laserdrucker) senden. Umfangreiche Informationsbestände können auf einem gemeinsamen Archiv-Server abgelegt werden.

Server = Bediener

Die Computersysteme, meist handelt es sich dabei um die am Arbeitsplatz verwendeten Computer, die Dienstleistungen von Servern in Anspruch nehmen, bezeichnet man als **Clients**. In manchen Fällen kann ein Computersystem gleichzeitig Client in Bezug auf eine Dienstleistung A und Server für andere Clients in Bezug auf eine Dienstleistung B sein.

Client = Klient,  
Kunde

### 1.1.6 Das Betriebssystem

Zur Steuerung und Verwaltung der einzelnen Komponenten eines Computersystems dient ein so genanntes Betriebssystem.

Das **Betriebssystem** erledigt unter anderem folgende Aufgaben:

Betriebssystem

- Voreinstellung der Hardware nach dem Einschalten auf definierte Anfangswerte (*bootstrapping*).
- Eröffnen des Dialogs zwischen Benutzer und Computer über Bildschirm, Tastatur und Maus (*console driver*).
- Interpretation und Ausführung der Kommandos des Benutzers an das Betriebssystem (*command interpreter*).
- Zuweisung verfügbarer Arbeitsspeicherbereiche an die einzelnen Programme (*memory manager*).
- Behandlung der Ein-/Ausgabe-Anforderungen (*I/O-manager*).
- Verwaltung der Peripherie (*resource manager*).
- Verwaltung der Dateien (*file manager*).

Das Betriebssystem selbst ist ein Programm wie jedes andere. Damit es jederzeit Kommandos vom Benutzer entgegennehmen kann, befindet sich ein Teil des Betriebssystems ständig im Arbeitsspeicher des Computers (speicherresidenter Teil). Nicht ständig benötigte Teile befinden sich auf einer so genannten Systemplatte, d.h. auf einem externen Speicher, und werden bei Bedarf in den Arbeitsspeicher geholt.

Es gibt unterschiedlich leistungsfähige und unterschiedlich stark verbreitete Betriebssysteme. Am weitesten verbreitet sind folgende:

- Windows 2000/XP von der Fa. Microsoft (für Clients),
- Windows NT/XP von der Fa. Microsoft (für Server und Clients),
- UNIX (für Server und Clients) und die kostenlose Variante Linux,
- MVS bzw. OS/390 (für IBM-Großcomputer).

Bevor Anwendungssoftware benutzt werden kann, muss sie auf dem entsprechenden Computersystem zunächst installiert werden (wenn sie noch nicht vorhanden ist). Abb. 1.1-7 beschreibt die dazu notwendigen Schritte sowie die Anwendung eines Software-Produkts.

### 1.1.7 Fallstudie: Die Firma ProfiSoft



Nach dem Studium der Informatik mit Schwerpunkt Software-Technik gründen die zwei ehemaligen Absolventen Alexander Klug, genannt Alex, und Robert Neumann, genannt Rob, die Firma ProfiSoft.

Ihre im Studium erworbenen Erkenntnisse wollen sie dazu nutzen, um für Kunden Software zu entwickeln, Kunden zu schulen und zu beraten.

Um bei Gesprächen mit potenziellen Kunden einen guten Eindruck zu hinterlassen, entwirft jeder ein persönliches Qualifikationsprofil. Abb. 1.1-8 zeigt das Profil von Alex. Es enthält zusätzlich eine gesprochene Begrüßung, falls das Profil als Datei an den Kunden gesandt wird.

## 1.2 Die Informatik

Parallel zur stürmischen Entwicklung der Computer- und Software-Technik entstand die Wissenschaftsdisziplin **Informatik**. Zunächst wurde die Informatik als Spezialgebiet innerhalb anderer wissenschaftlicher Disziplinen betrieben. Seit 1960 wird sie nicht mehr nur als eine Ansammlung von aus anderen Wissenschaften (z.B. Logik, Mathematik, Elektrotechnik) entliehenen Methoden und Regeln aufgefasst. Die Informatik hat sich vielmehr zu einem zusammenhängenden, theoretisch fundierten Gebäude, also zu einer neuen Grundlagenwissenschaft entwickelt, auf die andere Wissenschaften zurückgreifen.

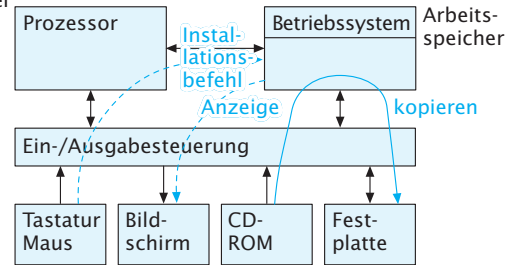
Die Informatik ist in mehrfacher Hinsicht eine grundsätzlich neuartige Wissenschaft /Brauer, Münch 96, S. 12 ff./:

- Die Hauptprodukte der Informatik sind immateriell, nämlich Software – im Unterschied zu den traditionellen Ingenieurwissenschaften.
- Die Produkte der Informatik sind im Allgemeinen erst in Verbindung mit materiellen Objekten praktisch nutzbar.
- Die Informatik ist potenzielle – und meist auch schon tatsächliche – Kooperationspartnerin für jede Wissenschaft und jede Sparte praktischer Tätigkeiten.

Obwohl diese Eigenarten der Informatik allgemein anerkannt sind, gibt es jedoch unterschiedliche Ansichten über das Selbstverständnis der Informatik. Es lassen sich folgende vier Positionen unterscheiden (Abb. 1.2-1):

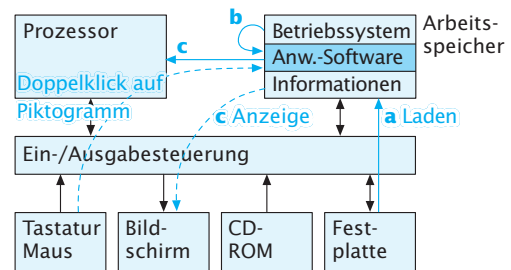
### Installation

- 1 Kauft man ein Software-Produkt, dann erhält man in der Regel ein Benutzerhandbuch sowie eine CD-ROM oder Disketten, die die Anwendungssoftware enthalten. (Auf die Bereitstellung über das Netz wird im Kapitel 2 eingegangen.)
- 2 Damit das Software-Produkt benutzt werden kann, muss die Software installiert werden. Installieren bedeutet in der Regel, dass die gesamte Software oder zumindest einige Teile vom Datenträger (Diskette oder CD-ROM) auf die Festplatte kopiert werden. Dies geschieht deshalb, weil die Software von der Festplatte schneller in den Arbeitsspeicher geladen werden kann, als von der CD-ROM oder der Diskette.
- 3 Das Kopieren übernimmt das Betriebssystem, d.h., es liest die Dateiinhalte vom Datenträger, legt Dateien auf der Festplatte an und trägt dort die Inhalte wieder ein.  
Die Anweisung zum Installieren muss vom Benutzer über die Tastatur/Maus an das Betriebssystem gegeben werden.  
In der Regel besitzt jede Anwendungssoftware ein Installationsprogramm, meist *setup* genannt, das durch Doppelklick mit der Maus auf dem Datenträger gestartet werden muss. Dieses Programm übernimmt dann unter Inanspruchnahme des Betriebssystems die Installation. Das Installationsprogramm erledigt folgende Aufgaben:
  - a Festlegen des Funktionsumfangs, der genutzt werden soll.
  - b Abstimmung der Software auf die vorhandene Hardware.
  - c Integration in bestehende Programme, z.B. um Daten auszutauschen.
  - d Anlegen eines anwendungsspezifischen Piktogramms auf der Arbeitsoberfläche (*desktop*), das nach dem Starten des Computersystems jeweils angezeigt wird.
- 4 Bei Lexika, Telefonbüchern usw. enthält die CD-ROM nicht nur die Software, sondern auch die Informationen, die verarbeitet werden sollen. In der Regel wird nur die Software auf die Festplatte kopiert, die Informationen bleiben auf der CD-ROM. Will man auf die Informationen zugreifen, muss die entsprechende CD-ROM in das Laufwerk eingelegt werden.



### Benutzung

- 5 Um Anwendungssoftware zu nutzen, muss sie gestartet werden. Dies geschieht in der Regel durch einen Doppelklick auf das entsprechende Piktogramm.  
In der Zentraleinheit läuft nun Folgendes ab:
  - a Das Betriebssystem lädt, d.h. kopiert, die Anwendungssoftware (ganz oder Teile davon) von der Festplatte in den Arbeitsspeicher.
  - b Anschließend startet das Betriebssystem die Anwendungssoftware.
  - c Die Anwendungssoftware wird vom Prozessor Schritt für Schritt abgearbeitet. Nach dem Laden wird die erste Anweisung ausgeführt. In der Regel wird das Anwendungsfenster auf dem Bildschirm angezeigt.
  - d Der Benutzer kann aus den angebotenen Funktionen die jeweils gewünschte aktivieren.



**Abb. 1.1-7: Installation und Benutzung von Anwendungssoftware**

- Position 1: Mathematisch-logische Orientierung,
- Position 2: Ingenieurwissenschaftliche Orientierung,
- Position 3: Evolutionäre Orientierung,
- Position 4: Partizipative Orientierung.

Entsprechend diesen Positionen wird der Begriff »Informatik« enger oder weiter gefasst, und die Schwerpunkte der Informatik werden verschieden gesehen.

Die »Gesellschaft für Informatik«, Interessenvertretung der deutschen Informatiker, verwendet folgende Definitionen:

Definitionen der Informatik  
www.gi-ev.de





### Profil von Dipl.-Inform. Alexander Klug

Gründer der Firma **ProfiSoft** zusammen mit Robert Neumann

#### Persönliche Angaben

- Geboren in Bochum
- Alter 24 Jahre

#### Studium

- Studium der Informatik
- Schwerpunkt: Software-Technik
- Diplomarbeit: Erstellung eines Java-Labors für die Informatikausbildung (mit Auszeichnung)
- Dauer: 10 Semester
- Während des Studiums Praktika in verschiedenen Softwarehäusern
- 4 Semester lang Hilfsassistent am Lehrstuhl für Software-Technik

#### Arbeitsschwerpunkte in der Firma ProfiSoft

- Innovative Seminare und Trainings auf dem Gebiet der Software-Technik
- Konzeption und Erstellung von multimedialen Firmenpräsentationen
- Konzeption und Erstellung von *e-learning*-Kursen

#### Hobbies

- Segeln
- Fischen



Abb. 1.1-8:  
Profil von Alex

»**Informatik** ist die Wissenschaft von der systematischen und automatisierten Verarbeitung von Information. Sie erforscht grundsätzliche Verfahrensweisen der Informationsverarbeitung und allgemeine Methoden ihrer Anwendung in den verschiedensten Bereichen. Für diese Aufgaben wendet die Informatik vorwiegend formale und ingenieurmäßig orientierte Techniken an. Durch Verfahren der Modellbildung sieht sie beispielsweise von den Besonderheiten spezieller Datenverarbeitungssysteme ab; sie entwickelt Standardlösungen für die Aufgaben der Praxis.« /GI 85/

»**Informatik** ist die Wissenschaft, Technik und Anwendung der maschinellen Verarbeitung und Übermittlung von Informationen.« /GI 87/

www.acm.org

Die ACM (Association for Computing Machinery) – größte internationale, wissenschaftliche Organisation der Informatiker – benutzt folgende Definition /Denning et al. 89/:



Computer Science

»**Computer science** and engineering is the systematic study of algorithmic processes – their theory, analysis, design, efficiency, implementation and application – that describe and transform information. The fundamental question underlying all of computing is, What can be (efficiently) automated?«

Informatik vs.  
Computer Science

Der Begriff *Computer Science* »schließt nämlich im Gegensatz zu Informatik u.a. den ganzen Bereich der betrieblichen Anwendungssysteme und ihres organisatorischen Umfeldes, also der Wirtschaftsinformatik und der ihr verwandten Gebiete, aus, die zumindest in den USA vorwiegend unter dem Oberbegriff *Information Systems* firmiert.« /Mayr, Maas 02, S. 178/

Im »Studien- und Forschungsführer Informatik« /Brauer, Münch 96/ wird Informatik aus einer anderen Perspektive betrachtet:

**Position 1: Mathematisch-logische Orientierung** /Dijkstra 89/

These: Computer können nur Symbole manipulieren. Sie tun dies mittels Programmen. Programme sind maschinell ausführbare Formeln. Aufgabe der Programmierer ist es, Formeln durch die Manipulation von Symbolen herzuleiten. Informatik befasst sich also mit dem Wechselspiel von maschineller und menschlicher Symbol-Manipulation. Die Informatik-Grundausbildung muss daher stringent mathematisch-logisch sein. Es ist – ohne Verwendung von Computern – die formale Manipulation einer einfachen imperativen Programmiersprache zu lehren. Ziel ist die Vermittlung der Fähigkeit, korrekte Programme zu schreiben, d.h. gegebene Spezifikationen korrekt in maschinell ausführbare Formeln umzusetzen.

**Position 2: Ingenieurwissenschaftliche Orientierung** /Parnas 90/

These: Informatiker arbeiten de facto wie Ingenieure, weil sie technische Artefakte herstellen. Es fehlt ihnen aber eine Ingenieursausbildung, die fundamentale Methoden wie z.B. Zuverlässigkeitsanalysen komplexer Systeme enthält. Es sollte daher eine Ingenieursausbildung angestrebt werden, die aus mathematischen und ingenieurwissenschaftlichen Kursen besteht. Die Programmierung realer Maschinen sollte im Grundstudium ignoriert werden.

**Position 3: Evolutionäre Orientierung** /Brooks 87/

These: Anforderungen an ein Software-System können bei realen Systemen nicht eindeutig formuliert werden. Spezifikationen erfassen häufig veraltete Anforderungen und verhindern so die rasche und flexible Anpassung an die sich ständig ändernden organisatorischen Strukturen und Aufgaben. Daher hat keine der bisher entwickelten formalen Methoden (algebraische Spezifikation, formale Verifikation, automatische Programmierung) zu einem qualitativen Sprung in der Software-Erstellung geführt. Kontinuierliche Verbesserungen sind die einzige Hoffnung.

**Position 4: Partizipative Orientierung** /Bonsiepen, Coy 92/

These: Computer sind Geräte, die bestimmte Funktionen innerhalb menschlicher Tätigkeitsbereiche erfüllen. Sie sind Mittel zum Zweck. Im Mittelpunkt stehen daher die alltäglichen Aufgaben von Menschen, die die Software nutzen. Software-Systeme sind Bestandteil einer funktionierenden Organisation. Der Informatiker muss deshalb den sozialen Kontext des Arbeitsplatzes, die Aufgabenverteilung zwischen Mensch und Computer innerhalb des Arbeitsprozesses verstehen lernen. Software-Systeme sind so zu entwickeln, dass sie einfach, benutzbar und arbeitsunterstützend sind. Ziel der Ausbildung muss es sein, Informatiker zu befähigen, Geräte, Programme und Prozesse zu spezifizieren und zu entwerfen, die den gestellten Anforderungen genügen. Es werden akzeptable Lösungen gesucht, nicht optimale und perfekte.

Quelle: /Bonsiepen, Coy 92/

**Abb. 1.2-1:**  
**Zum Selbstverständnis der Informatik**

Das Wort »Informatik« wurde 1968 vom damaligen Bundesforschungsminister Stoltenberg anlässlich der Eröffnung einer Tagung in Berlin verwendet. Mit Informatik wurde dabei der wissenschaftliche Hintergrund eines Forschungsprogrammes des Bundes auf dem Gebiet Datenverarbeitung bezeichnet. 1967 war in Frankreich das Wort »informatique« aufgekommen und hat sich inzwischen auch im Holländischen (informatika), Italienischen (informatica), Polnischen (informatyka), Tschechischen und Russischen eingebürgert. In den Vereinigten Staaten entstand Anfang der 60er-Jahre der Ausdruck »computer science«. Teilweise scheint sich sowohl im Angelsächsischen (informatics) als auch im internationalen Gebrauch der Wortstamm »Informatik« durchzusetzen.

»**Informatik** lässt sich also kennzeichnen durch die drei Begriffe Intelligenz – Formalismen – Technik oder als Intelligenzformalisierungstechnik.

Etwas allgemeiner könnte man sagen: Informatik ist die (Ingenieur-) Wissenschaft von der theoretischen Analyse und Konzeption, der organisatorischen und technischen Gestaltung sowie der konkreten Realisierung von (komplexen) Systemen aus miteinander und mit ihrer Umwelt kommunizierenden (in gewissem Maß intelligenten und autonomen) Agenten oder Akteuren, die als Unterstützungssysteme für den Menschen in unsere Zivilisation eingebettet werden müssen – mit Agenten/Akteuren sind Software-Module, Maschinen (zum Beispiel Staubsauger) oder roboterartige Geräte gemeint.«

/Langenheder, Müller, Schinzel 92/ charakterisieren **Informatik** als »Legierung aus Formalwissenschaft, Naturwissenschaft, Technik und Geisteswissenschaften«.

## LE 1 1 Einführung

/Freytag 93/ betrachtet **Informatik** ähnlich wie Architektur als Gestaltungswissenschaft (»Form und Kontext plus Methoden«).

Der geschichtliche Hintergrund der Informatik wird in Abb. 1.2-2 skizziert.

Informatik vs. IT

»Als grobe Abgrenzung der Informatik gegenüber der Informationstechnologie könnte man den stärker logisch abstrahierenden, softwareorientierten Blickwinkel der Informatik gegenüber dem mehr auf die technischen Basiskomponenten ausgerichteten, von Nachrichten- und Elektrotechnik bzw. Elektronik geprägten Herangehen der IT heranziehen.« /Mayr, Maas 02, S. 178/

Wegen des universellen Charakters der Informatik lässt sich ihr Gebiet schwer eingrenzen. Aufgaben, Gegenstände und Methoden der Informatik werden stark von den Natur-, Ingenieur- und Geisteswissenschaften beeinflusst. Die Informatik selbst liegt zwischen diesen Disziplinen. Von den Naturwissenschaften unterscheidet sich die Informatik, da ihr Forschungsgegenstand von Menschen geschaffene Systeme und Strukturen sind; von den Ingenieurwissenschaften unterscheidet sie sich ebenfalls, da ihre Betrachtungsgegenstände meist immateriell sind, und von den Geisteswissenschaften unterscheidet sich die Informatik, da sie sich nicht auf Erkenntnisgewinn und Beschreibung von Sachverhalten beschränkt, sondern praktisch anwendbare Ergebnisse erzielt. Zusammenfassend kann man sagen, dass die Informatik eine **Strukturwissenschaft** ist.

Dynamik

Dynamik ist für mich ein charakteristisches Merkmal der Informatik. Ohne das Verstehen der Abläufe in Programmen ist das Wesen und die Faszination der Informatik nicht zu begreifen. Dies erschwert aber auch dem Lernenden den Einstieg in die Informatik. Auf der beiliegenden CD-ROM werden daher wichtige Abläufe animiert dargestellt, um die Vorstellung zu unterstützen.



Gegenstandsbereich

Der Gegenstandsbereich der Informatik ist vielschichtig. Mindestens vier miteinander eng verzahnte Schichten gehören dazu /GI 87/:

- Hardware,
- Software,
- Organisationsstrukturen,
- Benutzer und Anwender.

Basierend auf einem eng gefassten Informatikbegriff definiert die »Gesellschaft für Informatik« /GI 85/ den Gegenstandsbereich folgendermaßen:

»Die Informatik befaßt sich daher

- a** mit den Strukturen, den Eigenschaften und den Beschreibungsmöglichkeiten von Information und Informationsverarbeitung,
- b** mit dem Aufbau, der Arbeitsweise und den Konstruktionsprinzipien von Rechnersystemen,





Die geschichtliche Entwicklung zeigt, dass es dem Menschen im Laufe der Zeit gelungen ist, sich seine Arbeit zu erleichtern. So entstanden von der einfachen Steinaxt über Handwerkzeuge aller Art die heutigen energieverarbeitenden Maschinen wie Motoren, Bagger, Fahrzeuge, Kräne. Diese Maschinen haben dem Menschen im physischen Bereich fast überall die routinemäßige Arbeit abgenommen. Die menschliche Tätigkeit kann sich auf die Planung und Überwachung konzentrieren.

Neben der Erleichterung körperlicher Arbeit war der Mensch von Anfang an auch bemüht, sich geistige Arbeit durch geeignete Hilfsmittel leichter zu machen (Abacus, Rechenschieber, Tabellen). Das Lösen von Problemen oder allgemeiner das Verarbeiten von Informationen bedeutet geistige Arbeit. Von einer gewissen Komplexität an lassen sich Probleme nur noch mit Hilfsmitteln lösen. Im Laufe der Wissenschaftsgeschichte sind die zu lösenden Probleme bzw. die zu verarbeitenden Informationen immer komplizierter geworden. Damit hat auch der zur Lösung bzw. Verarbeitung erforderliche Aufwand ständig zugenommen. Als Hilfsmittel entstanden vor etwa 50 Jahren elektronische Datenverarbeitungsanlagen, die als informationsverarbeitende Maschinen die Menschen bei der Lösung von Problemen unterstützten und es ihnen ermöglichen, ökonomisch und rationell zu arbeiten. Geistige Routinetätigkeiten werden uns von Computersystemen abgenommen.

Ein Computersystem kann Informationen wesentlich schneller verarbeiten als ein Mensch. Als geistiges Hilfsmittel des Menschen potenziert es damit seine Möglichkeiten und eröffnet ihm eine neue Dimension.

Durch die Geschwindigkeit heutiger Computer wird ein **Zeitraffereffekt** erzielt, durch den der Mensch in neue Bereiche eindringen kann und durch den neue Perspektiven sichtbar werden. Technische Spitzenleistungen wie der moderne Brückenbau, die Weltraumfahrt oder die Wahlhochrechnungen wären ohne ein Instrument wie ein Computersystem nicht möglich.

Die schnelle Verarbeitung von Informationen sowie die Aufbewahrung von umfangreichen Daten und die Auswertung dieser Daten nach unterschiedlichsten Kriterien potenziert auch Missbrauchsmöglichkeiten.

Es sollte daher jeder durch seine kritische Aufmerksamkeit dazu beitragen, dass die Annehmlichkeiten, die die automatische Informationsverarbeitung uns beschert, allen Menschen zugute kommen, dass die Missbrauchsmöglichkeiten dieser neuen Technik aber auf ein Minimum eingeschränkt werden.

**Abb. 1.2-2:  
Geschichtlicher  
Hintergrund  
der Informatik**

Abacus = auf Drähte aufgefädelt Perlen

- c** mit der Entwicklung sowohl experimenteller als auch produktorientierter informationsverarbeitender Systeme moderner Konzeption,
- d** mit den Möglichkeiten der Strukturierung, der Formalisierung und der Mathematisierung von Anwendungsgebieten in Form spezieller Modelle und Simulationen und
- e** mit der ingenieurmäßigen Entwicklung von Softwaresystemen für verschiedenste Anwendungsbereiche unter besonderer Berücksichtigung der hohen Anpassungsfähigkeit und der Mensch-Computer-Interaktion solcher Systeme.«

Die Informatik gliedert sich in die Teilbereiche (Abb. 1.2-3):

- Kerninformatik und
- Angewandte Informatik.

Die Kerninformatik beschäftigt sich mit den zentralen Forschungsgebieten der Informatik. Unter »Angewandter Informatik« versteht man Anwendungen von Methoden der Kerninformatik in anderen Fachwissenschaften.

## LE 1 1 Einführung

Schwerpunktbildungen innerhalb der Kerninformatik führen zu den Teilgebieten

- Theoretische Informatik,
- Praktische Informatik und
- Technische Informatik.

Die schnell wachsende Bedeutung der Informatik sowie die zunehmende Nachfrage nach ausgebildeten Fachleuten auf diesem Gebiet führte Ende der 60er-Jahre zur Schaffung des Studiengangs Informatik an Universitäten und Fachhochschulen.

Die »Rahmenordnung für die Diplomprüfung im Studiengang Informatik an Universitäten und gleichgestellten Hochschulen« /Rahmenordnung 95/ sagt Folgendes zum Informatikstudium:

Informatikstudium

Zitat

»Die Informatik ist zentral mit der Entwicklung und Beherrschung komplexer Informatik-Systeme befasst. Sie hat sich unbeschadet ihrer strukturwissenschaftlichen Grundlagen zu einer ingenieurmäßigen Disziplin im Sinne konstruierender Tätigkeiten entwickelt. Diese Entwicklung wird sich verstärkt fortsetzen.

Abb. 1.2-3:  
Gliederung der  
Informatik

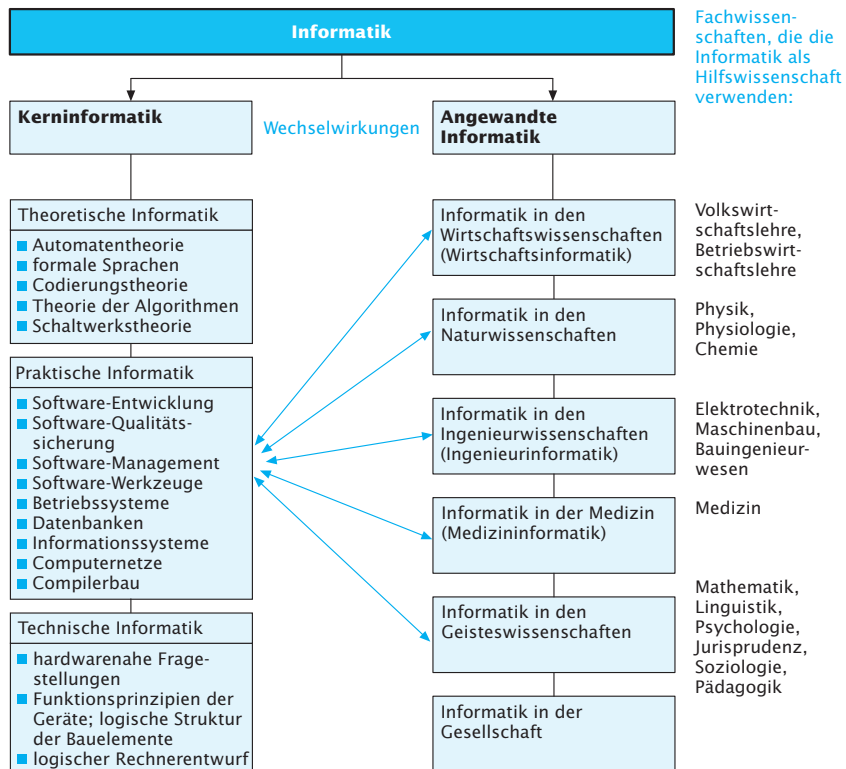
In Bezug auf Umfang und Struktur ist der Diplomstudiengang Informatik daher den Studiengängen in den klassischen Ingenieurwissenschaften gleichzusetzen.

Nachbar-  
wissenschaften  
(Schnittstellen)

Mathematik  
(mathem. Logik,  
Systemtheorie)

Elektrotechnik  
(Gerätetechno-  
logie, Nach-  
richtentechnik)

Physik  
(Halbleiter-  
technik, Fest-  
körperphysik)



Dem Diplom-Informatiker steht ein sehr vielfältiges Tätigkeitsfeld offen. Von ihm wird erwartet, dass er in der Lage ist, ingenieurmäßig Methoden und Techniken zur Konstruktion großer Informatik-Systeme anzuwenden, sich schnell in neue Problemstellungen und andere Gebiete einzuarbeiten und durch selbstständige Weiterbildung den Anschluss an die Entwicklung seines Faches zu halten. Dazu ist eine breite Ausbildung in den methodischen Grundlagen und die Vermittlung von ingenieurmäßigen Vorgehens- und Verhaltensweisen erforderlich.«

Die »Gesellschaft für Informatik« /GI 85/ zählt für das Grundstudium für die »Einführung in die Informatik« (ca. 22 Semesterwochenstunden) folgende Themen auf:

- Grundbegriffe,
- Methoden und Modelle der Informatik,
- Algorithmen,
- Konzepte der Programmierung und ihre Anwendung usw.



Im »Computing Curricula 91« der internationalen Informatikorganisationen ACM und IEEE (*Institute of Electrical and Electronics Engineers*) /Curricula 91/ werden für das Gebiet »Algorithmen und Datenstrukturen« (47 Vorlesungsstunden) folgende Themen vorgeschlagen:

[www.ieee.org](http://www.ieee.org)  
[www.computer.org](http://www.computer.org)

- Grundlegende Datenstrukturen
- Abstrakte Datentypen
- Rekursive Algorithmen
- Komplexitätsanalyse
- Komplexitätsklassen
- Sortieren und Suchen
- Berechenbarkeit und Unentscheidbarkeit
- Problemlösungsstrategien
- Parallele und verteilte Algorithmen

## 1.3 Gliederung und Aufbau dieses Buches

Die Curricula zur Grundlagenausbildung in der Informatik bieten dem Lehrenden genügend Spielraum für eigene inhaltliche Schwerpunkte und methodisch-didaktische Konzepte. Inhaltlich orientiert sich dieses Buch mehr an den in Abb. 1.2-1 dargestellten Positionen 3 und 4 als an den Positionen 1 und 2. Jahrelange Unterrichtserfahrung haben mir gezeigt, dass ein theoretischer, deduktiver Einstieg in die Grundlagen der Informatik *ohne* Computerbenutzung – wie es die Positionen 1 und 2 fordern – demotivierend und realitätsfern ist. Nur eigene Praxiserfahrungen mit Computersystemen von Anfang an geben dem Lernenden ein Gefühl für die Möglichkeiten der Programmierung, zeigen ihm aber auch deutlich die Probleme der Programmierpraxis – insbesondere die Probleme der Spezifikation und der Qualität. Mit

diesem Problembewusstsein ausgestattet, ist der Lernende aufnahmebereit für Themen wie Testen, Verifikation, Anforderungsmodellierung, Eigenschaften und Aufwand von Algorithmen.

induktiv vs.  
deduktiv

Die didaktisch-methodische Vorgehensweise orientiert sich daher vorwiegend am induktiven Vorgehen, d.h. vom Speziellen zum Allgemeinen. Ausgangspunkt können mehrere Beispiele sein, aus denen dann auf eine allgemeine Regel geschlossen wird. Diese Vorgehensweise wird aber nicht dogmatisch gesehen. Bei Bedarf wird auch deduktiv vorgegangen, d.h., aus dem Allgemeinen wird der Einzelfall bzw. das Besondere hergeleitet.

Lernen durch Beispiele, Analogie und eigenes Tun

»Lernen durch Beispiele«, »Lernen durch Analogie« und »Lernen durch eigenes Tun« fällt oft besonders leicht. Daher werden diese Lernformen – insbesondere durch die Aufgaben – unterstützt. Da man am meisten lernt, wenn man selbst etwas tut (Merkfähigkeit 90 Prozent), werden eigene Aktivitäten unterstützt. Zu vielen Lehreinheiten gibt es multimediale Unterstützungen und Ergänzungen, die auf der beigefügten CD-ROM bzw. über das Internet zur Verfügung stehen.

Bei jeder Einarbeitung in ein neues Gebiet muss man sich mit Prinzipien, Methoden, Konzepten, Notationen und Werkzeugen befassen.



**Prinzipien** sind Grundsätze, die man seinem Handeln zugrunde legt. **Methoden** sind planmäßig angewandte, begründete Vorgehensweisen zur Erreichung von festgelegten Zielen (im Allgemeinen im Rahmen festgelegter Prinzipien). Methoden enthalten also den Weg zu etwas hin, d.h., sie machen Prinzipien anwendbar. Methoden geben außerdem an, welche Konzepte wie und wann verwendet werden, um die festgelegten Ziele zu erreichen.

**Konzepte** erlauben es, definierte Sachverhalte unter einem oder mehreren Gesichtspunkten zu modellieren.

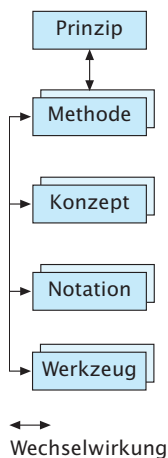
Eine **Notation** stellt Informationen durch Symbole dar. Ein Konzept kann durch eine oder mehrere Notationen dargestellt werden.

**Werkzeuge** (*tools*) dienen der automatisierten Unterstützung von Methoden, Konzepten und Notationen.

Zentraler Gegenstand der Informatik ist die Software. Eine Software-Entwicklung besteht aus drei Hauptaktivitäten:

- **Definition** der Anforderungen (Systemanalyse),
- **Entwurf** des Software-Systems (Programmieren im Großen),
- **Implementierung** bzw. **Programmierung** der Software-Komponenten (Programmieren im Kleinen).

Der inhaltliche Schwerpunkt dieses Buches liegt auf den Prinzipien, Methoden, Konzepten und Notationen der Programmierung von Software-Komponenten, auch »Programmieren im Kleinen« genannt. Für einen Informatiker sind solide Kenntnisse und Fertigkeiten auf diesem Gebiet unumgänglich, um die Möglichkeiten, Grenzen und Probleme der Realisierung von Software richtig einschätzen zu können.



Die Aktivitäten »Definition« und »Entwurf« werden ansatzweise bereits mitbetrachtet (siehe dazu auch: /GI 97/). Diese Thematik gehört aber in eine Vorlesung über Software-Technik (siehe z.B. »Lehrbuch der Software-Technik« /Balzert 98, 01/).

Der »rote Faden« dieses Buchs sind die Konzepte der Programmierung und Algorithmik. Die wichtigsten Konzepte der Software-Entwicklung und der Programmierung sind heute die Konzepte der »Objektorientierung«, wie Objekte, Klassen, Vererbung. Diese Konzepte werden in diesem Buch immer zuerst vorgestellt. Die entsprechenden Kapitelüberschriften lauten dazu: »Zuerst die Theorie: ...«. Anstelle von »Zuerst die Theorie« könnte auch stehen: »Zuerst die Konzepte«.

»roter Faden«



Konzepte werden durch eine oder mehrere Notationen beschrieben. Als grafische Notation für die Beschreibung objektorientierter Konzepte hat sich die UML (*unified modeling language*) durchgesetzt, die in diesem Buch eingeführt und benutzt wird.

UML

Zur Programmierung werden Programmiersprachen benutzt. Sowohl von den unterstützten Konzepten als auch von der Bedeutung her halte ich objektorientierte Programmiersprachen heute für die Grundlagenausbildung am geeignetsten. Andere Programmiersprachen wie deklarative, funktionale oder logikorientierte Sprachen sind in ihren Konzepten und Anwendungsbereichen zu spezialisiert, um als »erste« Sprache vermittelt zu werden.

Betrachtet man wichtige objektorientierte Sprachen wie Java, C++, C#, Smalltalk, Eiffel, Oberon, dann stellt sich die Frage, welche Sprache für die Ausbildung am besten geeignet ist.

Für mich sind folgende Kriterien am wichtigsten:

- In der Praxis weit verbreitet und eingesetzt.
- Unterstützung allgemein anerkannter Konzepte.
- Innovationsgrad der Sprache.
- Grad der Normung.
- Einfachheit und Eleganz der Sprache.
- Anwendungsspektrum der Sprache.

Eine Bewertung der Sprachen anhand dieser Kriterien führt zur Verwendung der Sprachen Java, C++ und C# in diesem Buch. Java steht als Sprache im Mittelpunkt, die speziellen C++- und C#-Konzepte werden ergänzend behandelt.

Die jeweils verwendete Programmiersprache beeinflusst natürlich stark die Reihenfolge der vorgestellten Konzepte und die Auswahl der Beispiele. Wegen der großen Bedeutung der Objektorientierung in der Software-Technik habe ich mich dazu entschlossen, von Anfang an die objektorientierte Sichtweise einzuführen und nicht zuerst die prozedurale Sicht zu verwenden. Bei jeder Einführung in ein Wissensgebiet sollte der Lernende nicht mit zu vielen Konzepten, Notationen und Begriffen auf einmal konfrontiert werden. Bei dem hier gewählten ganzheitlichen Einstieg in die Objektorientierung werden

am Anfang einige Konzepte nur erwähnt und bleiben zunächst im »Nebel«. Nach und nach »lichtet sich der Nebel«, wenn die erwähnten Konzepte detailliert betrachtet werden. Diesem Nachteil steht jedoch der große Vorteil gegenüber, dass von Anfang an die Kernkonzepte der Objektorientierung wie Objekte und Klassen im Mittelpunkt stehen und beherrscht werden.

Die Programmiersprachen-Notationen beginnen meistens mit Kapitelüberschriften wie »Dann die Praxis: ...«

**Vollständigkeit** Eine vollständige Darstellung der Grundlagen der Informatik ist heute in einem Lehrbuch *nicht* mehr möglich. Auf der einen Seite ist das Gebiet *nicht* eindeutig abgrenzbar, auf der anderen Seite ist die Innovationsgeschwindigkeit insbesondere bei den Programmiersprachen hoch. Daher erfordert eine Tätigkeit in der Informatik immer lebenslanges Lernen.

**Entdeckendes Lernen** Im vorliegenden Buch wird die Programmiersprache Java *nicht* vollständig behandelt. Ausgehend von exemplarischen Darstellungen sollen Sie, lieber Leser, dann durch »entdeckendes Lernen« weitere Sprachelemente selbstständig »erkunden«, sich die nötigen Informationen auf der beiliegenden CD-ROM oder im Internet besorgen und dadurch *die* Gebiete ihrem Wissen hinzufügen, an denen Sie Interesse haben.

**CASE** Um Software zu entwickeln, ist der Einsatz von Werkzeugen notwendig. Für die Programmierung sind Compiler und Programmierumgebungen unumgänglich. Der Begriff **CASE** steht für **Computer Aided Software Engineering**. Er drückt aus, dass Software-Entwicklung mithilfe von Software-Werkzeugen erfolgt. Man spricht daher auch oft von CASE-Werkzeugen, um zu betonen, dass der Einsatz von Software-Werkzeugen zum Zwecke der Software-Entwicklung gemeint ist.

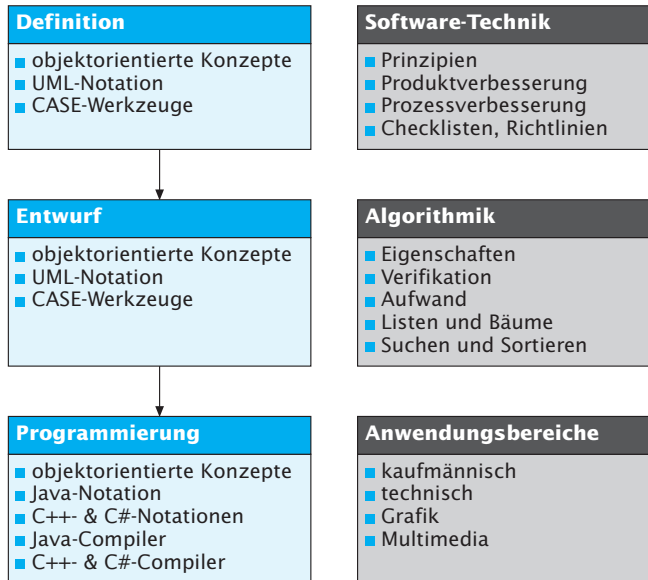
In diesem Buch wird davon ausgegangen, dass CASE-Werkzeuge eingesetzt werden. Auf der beigefügten CD-ROM befinden sich eine Reihe entsprechender Werkzeuge.



Die Struktur des Buches zeigt Abb. 1.3-1.

Globales Ziel des Buches ist es, einen systematischen Überblick über Prinzipien, Methoden, Konzepte und Notationen des »Programmierens im Kleinen« und seine Einordnung in die verschiedenen Kontexte zu geben. Dieses Wissen – verbunden mit den praktischen Übungen am Computersystem – soll den Leser befähigen,

- professionell effiziente Programme problemgerecht
- zu entwickeln,
- zu analysieren,
- zu überprüfen,
- adäquat in der UML zu beschreiben und
- in Java zu transformieren, zu übersetzen und auszuführen.

Abb. 1.3-1:  
Struktur des  
Buches

**Algorithmen** → Algorithmus.

**Algorithmus** (Plural: Algorithmen), Problemlösungsbeschreibung, die festlegt, wie ein Problem gelöst werden soll (→ Programm).

**Anwender** Mitglieder einer Institution oder Organisationseinheit, die zur Erfüllung ihrer fachlichen Aufgaben → Computersysteme einsetzen (→ Benutzer).

**Anwendungssoftware** → Software, die Aufgaben des → Anwenders mithilfe eines → Computersystems löst. Setzt in der Regel auf der → Systemsoftware der verwendeten → Hardware auf bzw. benutzt sie zur Erfüllung der eigenen Aufgaben.

**Arbeitsoberfläche** Teil einer → grafischen Benutzungsoberfläche, die dem Benutzer quasi als Ersatz für die Schreibtischoberfläche dient. Es können Objekte, Hilfsmittel, Geräte und Anwendungen auf ihr repräsentiert werden. Mithilfe der → direkten Manipulation können Funktionen ausgeführt werden. Über → Fenster erfolgt die Kommunikation des → Benutzers mit den Anwendungen.

**Arbeitsspeicher** Medium zur kurzfristigen Aufbewahrung nicht zu umfangreicher Information; Bestandteil der → Zentraleinheit; siehe auch → externer Speicher.

**Benutzer** Personen, die ein → Computersystem unmittelbar einsetzen und selbst bedienen (→ Anwender).

**Betriebssystem** Spezielles Programm eines → Computersystems, das alle Komponenten eines → Computersystems verwaltet und steuert sowie die Ausführung von Aufträgen veranlasst.

**Bildschirm** Ausgabegerät zum Anzeigen von Informationen.

**Bit (binary digit)** Binärzeichen, das nur jeweils einen von zwei Zuständen darstellen bzw. speichern kann, z.B. null oder eins. Acht Bits fasst man zu einem → Byte zusammen.

**Byte** Maßeinheit für die → Speicherkapazität. In einem Byte kann ein Zeichen gespeichert werden; siehe auch → Bit.

**Client** Vernetztes → Computersystem, das Dienstleistungen von → Servern in Anspruch nimmt.

**Computer** Technische Geräte, die umfangreiche Informationen mit hoher Zuverlässigkeit und großer Geschwindigkeit automatisch, gesteuert von → Programmen, verarbeiten und aufbewahren können; auch Rechner oder elektronische Datenverarbeitungsanlagen genannt.

**computer science** → Informatik.

**Computersystem** → Computer (→ Hardware) und → Programme (→ Software).



**Datei** Logisch zusammenhängender Informationsbestand (z.B. Kundenstammdatei), vergleichbar mit einer Kartei bei der manuellen Informationsverarbeitung.

**desktop** »elektronische« →Arbeitsoberfläche.

**Direkte Manipulation** Bedienungsform, bei der analog zu einem physikalischen Vorgang Objekte mit der →Maus auf der →Arbeitsoberfläche selektiert, bewegt und losgelassen werden (*pick, drag & drop*). In Abhängigkeit von der Zielposition können dadurch Funktionen wie Kreieren, Löschen, Kopieren, Drucken und Verschieben realisiert werden (generische Funktion).

**Externe Speicher** Speichermedien zur langfristigen Aufbewahrung von großen Informationsmengen (Diskettenspeicher, Plattenspeicher, CD-ROM-Speicher).

**Fenster** Rechteckiger Bereich auf dem Bildschirm, der von der →Anwendungssoftware zur Ein- und Ausgabe von Informationen und Kommandos benutzt wird.

**file** →Datei.

**Grafische Benutzungsoberfläche**

Grafikbildschirm bestehend aus einer →Arbeitsoberfläche und →Fenstern, über die der Benutzer mit der →Anwendungssoftware interagiert und kommuniziert.

**GUI** →grafische Benutzungsoberfläche.

**GUI-System** Software-System, das die →grafische Benutzungsoberfläche verwaltet und die Kommunikation mit der →Anwendungssoftware abwickelt.

**Hardware** Alle materiellen Teile eines →Computersystems.

**icon** →Piktogramm

**Informatik** Ingenieurwissenschaft von der theoretischen Analyse und Konzeption, der organisatorischen und technischen Gestaltung sowie der konkreten Realisierung von eigenständigen oder eingebetteten Software-Systemen.

**Maus** Kleines Kästchen, das mittels der Hand auf dem Schreibtisch bewegt wird. Entsprechend der Handbewegung bewegt sich der *Cursor* auf dem Bildschirm.

**Multimedia** Interaktive, medienintegrierte Software-Systeme, bei denen durch die zeitliche, räumliche und in-

haltliche Synchronisation unabhängiger Medien gewünschte Funktionen bereitgestellt werden. Es sollten mindestens drei Medien integriert sein, wobei zumindest ein Medium zeitabhängig sein sollte.

**Piktogramm** Grafisch abstrakte Darstellung von Objekten, Funktionen, Anwendungen, Geräten, Hilfsmitteln und Prozessen auf dem Bildschirm. Bei Anwendungssoftware im Bürobereich z.B. Ordner, Papierblatt, Papierkorb usw.; auch Ikone (*icon*) genannt.

**Programm** Streng formalisierter, eindeutiger und detaillierter →Algorithmus, der maschinell ausgeführt werden kann.

**Programmiersprache** Formalisierte Sprache zum Schreiben von →Algorithmen, die automatisch ausgeführt werden sollen.

**Prozessor** Der Teil der →Zentraleinheit, der Programmanweisungen aus dem →Arbeitsspeicher liest, ihre Ausführungen vornimmt, zu verarbeitende Informationen aus dem →Arbeitsspeicher liest sowie Zwischenergebnisse und Ergebnisse im →Arbeitsspeicher ablegt.

**RAM** *Random Access Memory*, d.h. Speicher mit direktem Zugriff auf alle Speicherzellen; andere Bezeichnung →Arbeitsspeicher (interner Speicher).

**Scanner** Abtastgerät, das es ermöglicht, Informationen auf Papier in Bitmuster umzuwandeln (→Bit). Dadurch ist es z.B. möglich, Fotos, Grafiken, handschriftliche Skizzen in elektronische Informationen zu wandeln.

**Server** Vernetztes →Computersystem, das →Clients Dienstleistungen zur Verfügung stellt.

**Software (SW)** →Programme, zugehörige Informationen und notwendige Dokumentation, die es zusammengefasst erlauben, mithilfe eines →Computersystems Aufgaben zu erledigen.

**Software-Produkt** Produkt, das aus →Software besteht.

**Software-System** System, dessen Systemkomponenten und Systemelemente aus →Software bestehen.

**Speicher** Medium zur Aufbewahrung von Informationen; →Arbeitsspeicher, →externer Speicher.

**Speicherkapazität** Umfang der Informationen, die in einem →Speicher aufbewahrt werden können; gemessen in KB, MB, GB (→Byte).

**Systemsoftware** →Software, die für eine spezielle →Hardware oder Hardwarefamilie entwickelt ist, um den Betrieb und die Wartung dieser Hardware zu ermöglichen sowie ihre funktionellen Fähigkeiten zu ergänzen.

**Tastatur** Eingabegerät zum manuellen Eintippen von Zeichen; besteht aus einem Eingabeteil für Zeichen, speziellen Sondertasten, *Cursortasten*, evtl. Zehnerblock und frei programmierbaren Funktionstasten.

**Zentraleinheit** Teil eines →Computers, in der die eigentliche Informationsverarbeitung stattfindet; besteht aus →Prozessor und →Arbeitsspeicher.



Ein **Computersystem** besteht aus dem materiellen **Computer (Hardware)** und den immateriellen **Programmen (Software)**. Der Computer selbst setzt sich aus der **Zentraleinheit** und der Peripherie (Eingabegeräte wie **Tastatur**, **Maus** und **Scanner**, Ausgabegeräte wie **Bildschirm** und Drucker, **externe Speicher**, Netzanschluss) zusammen. **Externe Speicher** dienen zur langfristigen Aufbewahrung von umfangreichen Informationsbeständen. Die Informationen werden in Form von **Dateien (files)** auf externen Speichern abgelegt. Die Größe eines Speichers wird als dessen **Speicherkapazität** bezeichnet. Die Maßeinheit für die Speicherkapazität ist das **Byte**, das wiederum aus acht **Bits** besteht.

Die **Zentraleinheit** besteht aus dem Prozessor, in dem die Programme Anweisung für Anweisung ausgeführt werden, und dem **Arbeitsspeicher (RAM)**, in dem Programme und Informationen, die zur momentanen Programmausführung benötigt werden, kurzfristig aufbewahrt werden. Externe Speicher und der Arbeitsspeicher bilden zusammen den **Speicher** des Computers.

**Programme** teilen dem Computer mit, welche Aufgaben auszuführen sind. Das **Betriebssystem**, selbst ein spezielles Programm, steuert und koordiniert das Computersystem mit seinen Komponenten. Programme werden in einer **Programmiersprache** formuliert, die festlegt, nach welchen Regeln (Syntax) die Programme geschrieben werden müssen und welche Bedeutung (Semantik) die einzelnen Programmkonstrukte haben.

Bevor ein Programm in einer Programmiersprache geschrieben wird, wird die allgemeine Problemlösung oft als **Algorithmus** – meist in verbaler Form – formuliert.

**Software**, auch **Software-System** oder **Software-Produkt** genannt, gliedert man in **Anwendungssoftware (application software)**, z.B. **Multimedia-Anwendungen**, und **Systemsoftware**. Beide bilden zusammen mit der Hardware ein Computersystem. **Benutzer** bedienen Computersysteme direkt, **Anwender** liefern Informationen für Computersysteme und nutzen ihre Ergebnisse.

Jeder Benutzer kommuniziert über eine elektronische **Arbeits-oberfläche (desktop)** mit dem Computersystem.

## LE 1 Zusammenhänge/Literatur

Die **grafische Benutzungsoberfläche (GUI)** besteht aus **Piktogrammen (icons)** und **Fenstern**. Mithilfe der **direkten Manipulation** können Objekte auf dem Bildschirm selektiert und Aktionen darauf angewandt werden. Die Verwaltung der Arbeitsoberfläche übernimmt ein **GUI-System**.

In Unternehmen und Verwaltungen sind Computersysteme in der Regel vernetzt (Intranet). Solche Netze können wiederum untereinander verbunden sein (Internet). Computersysteme, die in derartigen Netzen Dienstleistungen für andere Computersysteme zur Verfügung stellen, bezeichnet man als **Server**. Computersysteme, die solche Dienstleistungen in Anspruch nehmen, sind **Clients**.

Die Wissenschaft, die sich mit Computersystemen und insbesondere der Theorie und Praxis ihrer Software beschäftigt, bezeichnet man als **Informatik (computer science)**.

- Zitierte Literatur
- /Balzert 98/  
Balzert H., *Lehrbuch der Software-Technik*, Band 2, Heidelberg: Spektrum Akademischer Verlag 1998
  - /Balzert 01/  
Balzert H., *Lehrbuch der Software-Technik*, Band 1, 2. Auflage, Heidelberg: Spektrum Akademischer Verlag 2001
  - /BDB 03/  
Fast 30 Millionen Online-Konten in Deutschland, Bundesverband deutscher Banken, 2003, [www.bdb.de](http://www.bdb.de)
  - /BMFT 94/  
*Initiative zur Förderung der Software-Technologie in Wirtschaft, Wissenschaft und Technik*, Bundesministerium für Forschung und Technologie, Bonn, 1. August 1994
  - /Bonsiepen, Coy 92/  
Bonsiepen L., Coy W., *Eine Curriculardebatte*, in: *Informatik-Spektrum*, (1992) 15, S. 323–325
  - /Braucher, Münch 96/  
Braucher W., Münch S., *Studien- und Forschungsführer Informatik*, Berlin: Springer-Verlag 1996
  - /Brooks 87/  
Brooks F. P., *No Silver Bullet-Essence and Accidents of Software Engineering*, in: *IEEE Computer* 20, 1987, pp. 10–19
  - /Curricula 91/  
*Computing Curricula 1991*, in: *CACM*, June 1991, pp. 69–84
  - /Denning et al. 89/  
Denning P. J. et al., *Computing as a Discipline*, in: *CACM*, Jan. 1989, pp. 9–23
  - /Dijkstra 89/  
Dijkstra E. W., *On the Cruelty of Really Teaching Computer Science*, in: *CACM*, Dec. 1989
  - /Freytag 93/  
Freytag J., *Das Studium der Informatik an Fachhochschulen*, in: *Informatik als Schlüssel zur Qualifikation*, Berlin: Springer-Verlag 1993, S. 59–63
  - /GI 85/  
*Ausbildung von Diplom-Informatikern an wissenschaftlichen Hochschulen – Empfehlungen der Gesellschaft für Informatik*, in: *Informatik-Spektrum*, Juni 1985, S. 164–165



- /GI 87/  
*Aufgaben und Ziele der Informatik*, Arbeitspapier der Gesellschaft für Informatik, Bonn, 20. Okt. 1987
- /GI 97/  
 Ergänzende Empfehlungen der Gesellschaft für Informatik: *Lehrinhalte und Veranstaltungsformen im Informatikstudium*, in: Informatik-Spektrum, Oktober 1997, S. 302–306
- /IKT-Report 03/  
 IKT-Report, Zentrum für Europäische Wirtschaftsforschung, Mannheim, Juni 2003, [www.zew.de](http://www.zew.de)
- /Langenheder, Müller, Schinzel 92/  
 Langenheder W., Müller G., Schinzel B. (Hrsg.), *Informatik – cui bono?*, Berlin: Springer-Verlag 1992
- /Mayr, Maas 02/  
 Mayr H. C., Maas J., *Perspektiven der Informatik*, in: Informatik-Spektrum, 20.6.2002, S. 177–186
- /Necker 94/  
 Necker T., *Informations- und Kommunikationstechnik in Deutschland – Visionen und Realitäten*, in: Informatik-Spektrum, 1994, S. 339–341
- /Parnas 90/  
 Parnas D. C., *Education for Computer Professionals*, in: IEEE Computer 23 (1), 1990, pp. 17–22
- /Rahmenordnung 95/  
*Rahmenordnung für die Diplomprüfung im Studiengang Informatik an Universitäten und gleichgestellten Hochschulen*, 1995, in: /Brauer, Münch 96/
- /Schmid, Broy 00/  
 Schmid D., Broy M., ... *noch nicht zu spät*. Das Walberg-Memorandum zur Förderung der IT-Forschung, in: Informatik-Spektrum 23(2), S. 109–117, 2000
- /TNS Emnid 03/  
 Absolute Mehrheit: *Die Deutschen sind online*, (N)Onliner Atlas 2003, TNS Emnid, 2003, [www.tns-emnid.com](http://www.tns-emnid.com)

Alle Wissens- und Verstehens-Aufgaben befinden sich als *Multiple-Choice-Tests* in dem e-learning-Kurs zu diesem Buch. Diese Tests sollten Sie zunächst zufriedenstellend lösen, bevor Sie die folgenden analytischen und konstruktiven Aufgaben bearbeiten.



**1 Lernziel:** Anhand von Beispielen Übertragungszeiten bei vernetzten Computersystemen berechnen können.

**a** Gegeben seien folgende Dateien eines Software-Projektes:

Index.html	10 kB
Bild1.gif	100 kB
Bild2.jpg	30 kB
Musik.au	25 kB

Wie viel Zeit benötigen ein Modem mit 14.400 Bits pro Sekunde und eine ISDN-Karte eines PC mit 64.000 Bits/Sekunde mindestens, um das Software-Projekt von einem Server auf einen Client zu übertragen?

**b** Wie viel Zeit benötigen Sie zusätzlich, wenn Sie ein eingescanntes Bild aus Ihrem letzten Urlaub mit der Größe 989 KB einfügen?

Analytische  
 Muss-Aufgabe  
 20 Minuten

## LE 1 Aufgaben

Konstruktive  
Muss-Aufgabe  
30 Minuten

**2** *Lernziel: Software installieren und benutzen können.*

- a** Verfassen Sie einen kurzen Text, in dem Sie sich vorstellen. Verwenden Sie ein Textverarbeitungssystem Ihrer Wahl. Orientieren Sie sich an der Fallstudie von Abschnitt 1.1.7.
- b** Scannen Sie ein Sie darstellendes Foto ein und platzieren Sie es in den Text aus **a**.
- c** Sprechen Sie einen Begrüßungstext und nehmen Sie ihn mit einer geeigneten Software auf. Integrieren Sie ihn in den Text aus **a**.

# 1 Einführung – Internet, Web und HTML



- Aufbau, Technik, Anschluss- und Adressierungsmöglichkeiten des Internet erläutern können.
- Die besprochenen Dienste mit ihren Charakteristika beschreiben können.
- Multimediale XHTML-Dokumente bezogen auf die Syntax analysieren können.
- Mithilfe der angegebenen Befehle multimediale XHTML-Dokumente erstellen und auf Web-Servern bereitstellen können.
- Anhand der beschriebenen Suchstrategien systematisch im Internet unter Einsatz von Suchmaschinen Informationen recherchieren können.
- Einen Web-Browser so bedienen können, dass die besprochenen Dienste im Internet in Anspruch genommen werden können.
- Einen Web-Browser um *plug-ins* erweitern können.

verstehen

anwenden



- Das Kapitel 1.1 muss bekannt sein.



- 1.4 Das Internet 36
  - 1.4.1 Der Aufbau des Internet 36
  - 1.4.2 Der Anschluss an das Internet 38
  - 1.4.3 Die Adressierung im Internet 39
- 1.5 Dienste im Internet 40
  - 1.5.1 Elektronische Post im Internet (E-Mail) 41
  - 1.5.2 Nachrichtengruppen im Internet (*newsgroups*) 45
  - 1.5.3 Plaudern im Internet (*chat*) 48
  - 1.5.4 Dateien übertragen im Internet (*ftp*) 49
  - 1.5.5 Computersysteme fernbedienen über das Internet (*Telnet*) 50
- 1.6 Das *World Wide Web* (WWW) 50
  - 1.6.1 Web-Browser 52
  - 1.6.2 Suchen und Finden im Internet 54
- 1.7 Die Sprachen HTML und XHTML 57
  - 1.7.1 Grafiken und Bilder einbinden 60
  - 1.7.2 Erweiterungen von Web-Browsern (*plug-ins*) 62
  - 1.7.3 Erstellen von XHTML-Seiten und Bereitstellen auf Web-Servern 64
- 1.8 Fallstudie: Die Firma ProfiSoft im Internet 65

## 1.4 Das Internet

Computersysteme können für sich isoliert eingesetzt oder vernetzt mit anderen Computersystemen betrieben werden. Sind Computersysteme miteinander vernetzt, d.h. durch Kabel oder Funk miteinander verbunden, dann können Informationen zwischen diesen Computersystemen ausgetauscht und gegenseitig Dienstleistungen in Anspruch genommen werden. Das bekannteste und am meisten genutzte weltweite Netz, das Computersysteme verbindet, ist das Internet.

Definition Internet

Das **Internet** besteht aus

- einer Vielzahl von Computern,
- die direkt oder indirekt miteinander verbunden sind,
- die dasselbe Übertragungs-Protokoll (TCP/IP) verwenden,
- auf denen Dienste angeboten und/oder genutzt werden,
- einer Vielzahl von Benutzern, die von ihrem beruflichen oder privaten Computer-Arbeitsplatz aus direkten Zugriff auf diese Dienste haben,
- einer Vielzahl weiterer Netze, die über Kommunikationscomputer (*gateways*) erreichbar sind /Schneider 95/.

Das Internet stellt somit eine Kommunikationsinfrastruktur zum gleichberechtigten Informationsaustausch zur Verfügung, analog zum Telefonnetz, das eine Infrastruktur für die Sprachkommunikation bietet.

### 1.4.1 Der Aufbau des Internet

Das Internet ist durch *keine* zentrale Organisation aufgebaut worden. Vielmehr schließen sich Organisationen, z.B. Universitäten, zusammen und mieten Leitungen zwischen den einzelnen Mitgliedern mit kostenlosem oder pauschal mitfinanziertem Nutzungsrecht *aller* Teilstrecken für *alle* Mitglieder.

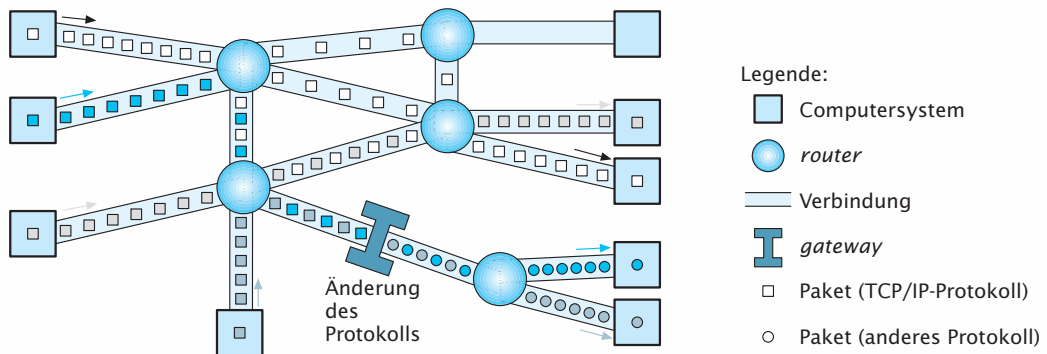
**Beispiel** Gibt es zwischen A und B sowie zwischen B und C eine Verbindung, dann kann nicht nur B mit A und C kommunizieren, sondern auch A mit C über den gemeinsamen Partner B.

Diese Art der Netznutzung ist konträr zum Telefonieren, das unmittelbar auf dem Verursacherprinzip beruht.

TCP/IP Die Kommunikation zwischen den Computersystemen im Internet geschieht durch das Übertragungsprotokoll **TCP/IP** (*transmission control protocol/internet protocol*). Die zu übermittelnden Datenströme werden bei diesem Übertragungsprotokoll in Pakete einheitlicher Größe (siehe Abb. 1.1-7) aufgeteilt, die voneinander unabhängig auf verschiedenen Wegen zu unterschiedlichen Zeiten zum Ziel kommen können. Jedes Paket wird mit Kopfdaten (*header*) versehen, die Angaben über die Absender- und Zielcomputer ent-



halten. Spezielle Wegplanungscomputer in den Netzknoten, *router* genannt, sind für die Auswahl der Teilstrecken verantwortlich. Sie analysieren die in einem eintreffenden Datenpaket gespeicherte Zieladresse und ermitteln aufgrund ihrer internen Adresstabellen (*routing tables*) den weiteren Weg des Datenpakets durch das Netzwerk. Beim Ausfall einer Übertragungsstrecke können sofort alternative Wege geschaltet werden, sodass Netzzusammenbrüche weitgehend vermieden werden. Im Zielcomputer werden die einzelnen Datenpakete wieder zu einem vollständigen Datenstrom in der richtigen Reihenfolge zusammengesetzt. Verloren gegangene Pakete werden automatisch wieder angefordert (Abb. 1.4-1).



Da das Übertragungsprotokoll TCP/IP die Datenströme in Pakete aufteilt, kann eine Leitung durch viele Computersysteme gemeinsam genutzt werden. Eine durchgehende Leitung zwischen Sender und Empfänger ist nicht nötig. Dadurch wird die Datenübertragung wesentlich effizienter. Da jeder Benutzer die Ressourcen anderer Partner mitbenutzt, sollte er die Dienste sorgfältig und bewusst nutzen. Wegen der Pauschalfinanzierung meinen neue Benutzer meist, dass die Benutzung des Internet nichts kostet.


Nicht alle Computernetze benutzen das Übertragungsprotokoll TCP/IP. Damit Computersysteme solcher Netze ebenfalls mit dem Internet kommunizieren können, müssen Kommunikationscomputer, so genannte **gateways** (Torweg, Einfahrt), zwischengeschaltet werden. Diese isolieren die beteiligten Netze total voneinander. Es wird keinerlei Protokollinformation des einen Netzes an das andere Netz weitergegeben, sondern es werden nur die Nutzdaten übergeben. Die Nutzdaten werden aus dem Protokollrahmen des Quellnetzes extrahiert und anschließend in die entsprechenden Datenformate des Zielnetzes umgesetzt.

Durch die inzwischen weite Verbreitung des Internet wird die Technik des Internet zunehmend auch für firmeninterne Netze verwendet.

Abb. 1.4-1:  
Datentransport im  
Internet

gateway

## LE 2 1.4 Das Internet

- Intranet** Ein **Intranet** verbindet Computersysteme eines Unternehmens oder einer Organisation basierend auf der Technik des Internets, insbesondere des TCP/IP-Übertragungsprotokolls, miteinander, ist aber *kein* Teil des öffentlichen Internets.
- Extranet** Ein **Extranet** erweitert ein Intranet zu anderen Unternehmen hin, etwa zu Händlern, Distributoren und Lieferanten. Es gibt einer definierten Benutzergruppe einen beschränkten Zugriff auf Informationen, die für sie relevant sind.
- zur Historie Historisch betrachtet liegen die Ursprünge des Internet in den USA. Dort forderten in den sechziger Jahren Universitäten und Forschungseinrichtungen ein landesweites Computernetzwerk. 1972 wurde daraufhin von der ARPA (*advanced research project agency*), einer Behörde der US-Regierung, das ARPAnet gegründet, das ursprünglich vier Computersysteme von vier Universitäten verband. 1974 wurde in einem Artikel von V. Cerf und R. Kahn die TCP/IP-Architektur in Grundzügen beschrieben. Im Anschluss hieran wurde das TCP/IP-Übertragungsprotokoll nach und nach entwickelt und im ARPAnet als alleiniges Protokoll verwendet. Wegen der Robustheit gegenüber Störungen und Ausfällen übernahm das US-Militär diese Technik und erstellte das DARPA-Netz (*Defense ARPA*), das über verschiedene Zwischenstufen zum heutigen Internet führte. 

### 1.4.2 Der Anschluss an das Internet

Lokale Computernetze können über Standleitungen, Wählleitungen oder paketvermittelnde Netze an das Internet angeschlossen werden.

Der private Benutzer kann sich über das analoge Telefonnetz mit einem Modem, über das digitale Telefonnetz mit einer ISDN-Karte oder über ein DSL-Modem in das Internet einwählen. Technisch ist dazu ein Programm nötig, das das TCP/IP-Übertragungsprotokoll für eine solche Punkt-zu-Punkt-Verbindung bereitstellt. Häufig verwendete Protokolle für diesen Zweck sind SLIP (*serial line internet protocol*) und PPP (*point-to-point-protocol*), die auf TCP/IP aufsetzen. Der Unterschied der beiden Protokolle besteht in der unterschiedlichen Fehlererkennung und Fehlerkorrektur. Sicherer ist das PPP.

Als Student oder Mitarbeiter einer Hochschule hat man im Allgemeinen die Möglichkeit, sich vom heimischen Computer aus über die Hochschule direkt ins Internet einzuwählen (siehe nächsten Abschnitt).

Als Firma oder Privatperson muss man einen Internet-Dienstanbieter, einen so genannten *Internet Service Provider* (ISP) – kurz Provider genannt – in Anspruch nehmen.

*provide* = bereitstellen, beliefern

Ein **Provider** ist eine Institution oder ein Unternehmen, die ein an das Internet angeschlossenes Netzwerk für Wählzugänge bereitstellt.

### 1.4.3 Die Adressierung im Internet

Jedes Computersystem im Internet besitzt eine eindeutige »Rufnummer«, seine **IP-Adresse**. Sie besteht aus vier Bytes. Diese vier Bytes werden dezimal, durch Punkte getrennt, geschrieben.

IP = internet  
protocol

Die IP-Adresse der Ruhr-Universität Bochum lautet 134.147.x.x. Der Lehrstuhl für Software-Technik hat die IP-Adresse 134.147.80.x.

Die weltweit eindeutige Kennzeichnung der Computersysteme im Internet übernehmen die NICs (*Network Information Centers*). Das deutsche Internet wird von DE-NIC verwaltet, das vom Rechenzentrum der Universität Karlsruhe betrieben und von den deutschen Internet-Providern gemeinsam finanziert wird.

NIC

Jeder Betreiber eines lokalen Netzes erhält vom gewählten Internet-Provider einen Block zusammenhängender Adressen und verteilt diese an die angeschlossenen Computersysteme.

Es gibt zurzeit drei Klassen von Adressen, die der Größe der jeweiligen Einrichtung entsprechen:

Adress-Klassen

- **Class-A-Adresse:** Aufbau A.x.x.x. mit  $0 \leq A \leq 126$

Auf der Welt gibt es genau 126 *Class-A*-Adressen (10.x.x.x wird nicht vergeben).

- **Class-B-Adresse:** Aufbau B1.B2.x.x. mit  $128 \leq B1 \leq 191$ ,  $0 \leq B2 \leq 255$

Mit dieser Adresse kann man ca. 16.000 Computer verwalten. 55 Prozent aller *Class-B*-Adressen sind bereits verteilt.

- **Class-C-Adresse:** Aufbau C1.C2.C3.x mit  $192 \leq C1 \leq 255$ ,  $0 \leq C2$ ,  $C3 \leq 223$

Sie erlaubt bis zu 255 Computer in einem Adressblock. Von *Class-C*-Adressen sind erst 4 Prozent vergeben.

Die Blockung der Adressen erleichtert die Netzverwaltung. Die Ruhr-Universität Bochum hat die *Class-B*-Adresse 134.147.x.x. Statt nun alle Computer der Universität in internationalen Tabellen einzutragen, genügt der Vermerk, wie der zentrale Zugangspunkt (*router*) der Universität zu erreichen ist. Alle Universitätscomputer befinden sich aus Netzsicht hinter diesem Zugangspunkt.

Da die IP-Adressen schlecht zu merken sind, erhalten die Computer im Internet zusätzlich einen oder mehrere funktionsbezogene Namen wie z.B. swt.ruhr-uni-bochum.de.

funktions-  
bezogene Namen

Diese **domain-Namen**, auch **DNS** (*domain name system*) genannt, sind leichter zu behalten und erleichtern es, ein Computersystem zu lokalisieren.

Sie sind in der Regel folgendermaßen aufgebaut:  
computer.bereich.institution.land

Die letzte Silbe im DNS bezeichnet man als *top level domain* (TLD). Sie gibt in der Regel das Land an, wobei insbesondere in den USA eine weitere Gliederung verwendet wird:

DNS

## 1.5 Dienste im Internet

- |                      |                                     |
|----------------------|-------------------------------------|
| ■ de für Deutschland | ■ edu Computer im US-Bildungssystem |
| ■ fr für Frankreich  | ■ com Computer in US-Firmen         |
| ■ ch für Schweiz     | ■ gov Computer in US-Verwaltungen   |
| ■ at für Österreich  | ■ org Computer in US-Organisationen |

Während die Endungen .edu und .gov nur für Bildungseinrichtungen in den USA bzw. Regierungsbehörden in den USA zugelassen sind, können die Endungen .com und .org von jedem benutzt werden.

neue  
top level domains Seit 2001 gibt es folgende neue *top level domains*, die unabhängig von Ländern sind:

- name für die Verwendung von Privatpersonen,
- prof für bestimmte Berufsgruppen oder *Professionals*,
- museum ausschließlich für Museen,
- biz für *business sites* mit rein kommerziellem Charakter,
- aero für die Luftfahrt,
- info für Informationsanbieter,
- coop für genossenschaftlich organisierte und betriebene Unternehmen und Organisationen.

Das Kürzel vor der *top level domain* soll die jeweilige Institution möglichst gut benennen.

Diese Namen werden von den NICs vergeben, wobei Benutzerwünsche berücksichtigt werden. Ist der gewünschte Name aber bereits belegt, dann muss im Allgemeinen ein anderer gewählt werden. Ausnahmen gibt es z.B. bei Städte- und Markennamen. Die weitere Systematik bleibt der jeweiligen Institution überlassen.

Nameserver Wird eine DNS-Adresse zur Adressierung eines Zielcomputers verwendet, dann werden die Daten zuerst an einen *Nameserver* geschickt. Ein *Nameserver* ist ein Computersystem, in dem DNS-Adressen in Tabellen gespeichert sind und das DNS-Adressen in IP-Adressen übersetzt.

Durch die DNS-Adressen ist es möglich, dass Adressaten im Internet zwischen Computersystemen »umziehen«, ohne dass die »Sender« informiert werden müssen. Das zuständige NIC überwacht und koordiniert die *Nameserver*-Funktionen der einzelnen Einrichtungen und betreibt den nationalen *Nameserver*, der die erste Anlaufstelle für die Anfragen aus dem Ausland darstellt. Diese Server werden *root*-Server genannt.

## 1.5 Dienste im Internet

Das Internet ist deshalb so interessant, weil es aus vielen verschiedenen Diensten besteht. Die wichtigsten Dienste sind:

- Elektronische Post im Internet (E-Mail),
- Nachrichtengruppen (*newsgroups*),
- Plaudern im Internet (*chat*),
- Dateien übertragen im Internet (*ftp*),

### 1.5.1 Elektronische Post im Internet (E-Mail)

- Computersysteme fernbedienen über das Internet (*telnet*),
- WWW (*world wide web*), kurz Web genannt.

Alle diese Dienste benutzen eigene – auf IP aufsetzende – Protokolle, um ihre Dienstleistungen zu erbringen.

Im Folgenden werden die einzelnen Dienste kurz vorgestellt. Wegen der Bedeutung des Web wird dieses in einem eigenen Kapitel behandelt.

#### 1.5.1 Elektronische Post im Internet (E-Mail)

Die elektronische Post (**E-Mail**, *electronic mail*) ist der am häufigsten genutzte Internet-Dienst. Manchmal in Sekunden, aber fast immer innerhalb eines Tages, erreicht eine E-Mail den Empfänger, egal wo sich sein Computer befindet. E-Mail

Für die Übertragung von E-Mails wurde ein spezielles Protokoll, das SMTP (*simple mail transport protocol*), entwickelt. Ein Computersystem, über das die Zustellung der elektronischen Post erfolgt, bezeichnet man dementsprechend als SMTP-Server.

Damit man elektronische Post empfangen kann, benötigt man eine E-Mail-Adresse. Technisch gesehen bedeutet dies, dass auf dem SMTP-Server des Providers ein elektronisches Postfach (*mailbox*) existiert. Die für diese Adresse eintreffende Post wird in diesem Postfach – es handelt sich dabei um eine Datei – gespeichert. Um die eingegangene Post zu lesen, müssen mit einer E-Mail-Software die Verbindung zu dieser Datei hergestellt und die Daten in das eigene Computersystem übertragen werden.

Eine E-Mail-Adresse ist folgendermaßen aufgebaut:

E-Mail-Adresse

Empfänger@smtp.server

wobei smtp.server in der Regel die DNS-Adresse ist und Empfänger meist der Vor- und Nachname oder das Monogramm des Adressaten ist, z.B.

hb@swt.ruhr-uni-bochum.de

Beide Angaben werden durch das @-Zeichen (sprich »Klammeraffe« oder »at« (engl. bei)) getrennt. Die *router* im Internet bearbeiten nur den rechten Teil der Adresse, der empfangende SMTP-Server nur den linken Teil. @-Zeichen

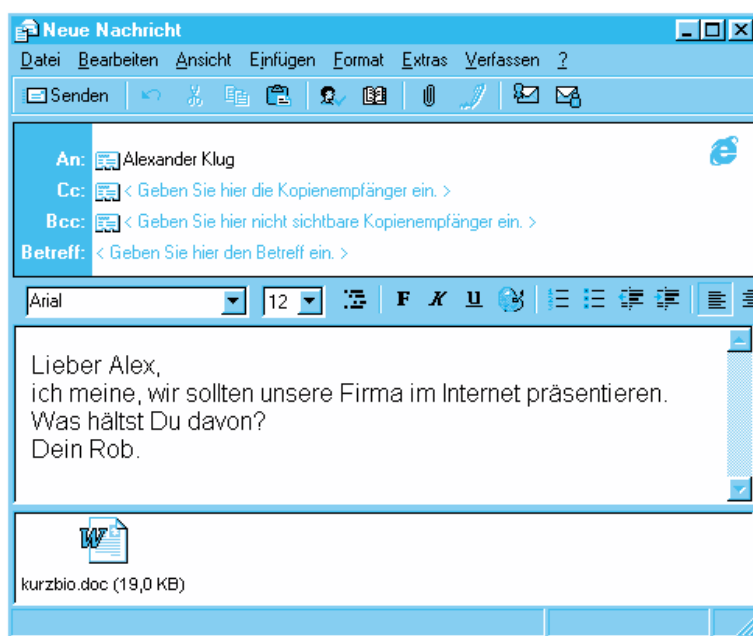
Eine E-Mail wird technisch gesehen in zwei Schritten befördert. Das Internet transportiert die elektronische Post bis zum gewünschten SMTP-Server. Anschließend speichert der Server die Post in das Postfach des Empfängers.

Jede E-Mail besteht aus einem Briefkopf (*header*) und der Nachricht selbst.

Abb.1.5-1 zeigt an einem Beispiel den Aufbau einer »elektronischen« Nachricht. Beispiel

Abb. 1.5-1: Beispiel  
für ein E-Mail-  
Programm

Beispiel



Das @-Zeichen wurde von dem Programmierer Ray Tomlinson 1972 für den Aufbau einer E-Mail-Adresse verwendet. Er suchte nach einem Zeichen, das niemals im Namen eines Menschen auftauchen würde und damit als eindeutiger Trenner dienen konnte. Auf der Tastatur, die er verwendete, ein Modell 33 Teletype, fand er das Zeichen @, das alle Vorgaben erfüllte. Das @-Zeichen selbst soll eine Entsprechung des französischen à sein. Englische Kaufleute sollen das früher so auf ihre Preisschilder geschrieben haben. Der Klammersatz heißt daher in der englischsprachigen Welt auch *commercial a*. Deshalb ist er bereits auf den ersten amerikanischen Schreibmaschinentastaturen zu finden /Kühnert 97/.

Der Briefkopf enthält folgende Angaben:

- An (To):  
Gibt den oder die Empfänger der Nachricht an.
- CC:  
Elektronische Post hat den Vorteil, dass sie ohne Aufwand beliebig vervielfältigt werden kann. Daher ist es einfach, sie nicht nur an eine Person, sondern an viele zu versenden. Weitere Empfänger können entweder hinter To angegeben werden, oder, wenn sie die Nachricht nur zur Kenntnis erhalten sollen, hinter CC (steht für *carbon copy*, der englische Begriff für Durchschlag).
- BCC:  
Sollen die angeschriebenen Personen nicht erfahren, dass mehrere Empfänger existieren, dann verwendet man BCC (*blind carbon copy*).
- Betreff (Subject):  
Zunächst sieht der Empfänger nur diesen Betreff und kann entscheiden, ob er die Post lesen will. Der Betreff steht sozusagen auf dem Briefumschlag.
- Anlage (Attachment):  
Gibt dem Empfänger an, dass die elektronische Nachricht ein oder mehrere Anlagen enthält. Beim Senden der Nachricht gibt der Autor an, welche Datei oder Dateien der Nachricht beigelegt werden sollen. In Abb. 1.5-1 ist die Anlage in Form eines Piktogramms im unteren Teil des Fensters dargestellt.

E-Mail wurde ursprünglich nur für Textmitteilungen in ASCII-Format konzipiert. Der ASCII-Code (*American Standard Code for Information*

**ASCII** (USASCII) 1963 genormter amerikanischer 7-Bit-Zeichensatz für den Fernschreibverkehr und den Datenaustausch zwischen Computern. Von den 128 Positionen benutzt er nur 100; er hatte nur Großbuchstaben. 1965 beschloss die ECMA eine Normzeichentabelle, ECMA-6. Sie entsprach im Wesentlichen ASCII, fügte aber auf den unbesetzten Positionen die 26 Kleinbuchstaben hinzu. 10 Positionen sollten nationalen Sonderzeichen zur Verfügung stehen. Diese Tabelle wurde 1968 vom ANSI als X3.4 übernommen und zum amerikanischen Normzeichensatz. 1974 wurde sie internationale Norm (ISO 646, revidiert 1983 und 1991) und gleichzeitig deutsche Norm (DIN 66003).

**Latin** Durch Hinzunahme eines achten Bits zu einem 7-Bit-Code wie ASCII ergeben sich 128 weitere Positionen in der Zeichentabelle. Diese werden heute auf verschiedene Weise genutzt. 1981 führte IBM für DOS-Computersysteme die *Codepage 437* ein, die die hinzugewonnenen Plätze für den amerikanischen Datenverkehr nutzte, aber auch die großen westeuropäischen Verkehrssprachen berücksichtigte. Im März 1985 beschloss ECMA einen 8-Bit-Code mit 256 Zeichen, der alle westeuropäischen Sprachen abdeckte (ECMA-94). Er wurde 1986 von der ISO zur Weltnorm gemacht (ISO 8859-1, genannt **Latin-1**) und ab 1987 nach und nach durch neun weitere Zeichensätze, zum Teil auch für nicht lateinische Alphabete, ergänzt (ISO 8859-1 bis 10) – die Nummer 2 (Latin-2) deckt die mittelosteuropäischen ab. Die IBM-*Codepage 819* für DOS ist identisch mit Latin-1. Meist wird mit DOS dagegen *Codepage 850* verwendet, die sich mit Latin-1 im Zeichenvorrat deckt, die Zeichen aber in anderer Reihenfolge anordnet. »Windows« enthält sämtliche Zeichen aus Latin-1 und noch einige mehr. Die deutsche Entsprechung zu ISO 8859-1 ist DIN 66303 (November 1986).

**Unicode** (UCS) Seit 1987 ursprünglich von den Firmen Apple und Xerox entwickelter Code, der die Schriftzeichen aller Verkehrssprachen der Welt aufnehmen soll: alphabetische wie syllabische und logografische Schriften. Er beruht auf 16 Bit (2 Byte) und stellt damit 65.469 Positionen zur Verfügung. Bisher wird er nur in Windows NT und der Programmiersprache Java verwendet. Im Juni 1992 wurde er zur internationalen Norm (ISO/IEC 10646-1).

Linke Tabellenhälfte: ASCII – Die darstellbaren Zeichen aus der Zeichentabelle

»ASCII« (ISO 646).

Beide Hälften: Latin-1 – Die darstellbaren Zeichen aus der Zeichentabelle »Latin-1« (ISO 8859-1).

SP	0	@	P	`	p		°	À	Ð	à	ä
!	1	A	Q	a	q		±	Á	Ñ	á	å
"	2	B	R	b	r		²	Â	Ò	â	ö
#	3	C	S	c	s		³	Ã	Ó	ã	ó
\$	4	D	T	d	t		´	Ä	Ô	ä	ô
%	5	E	U	e	u		µ	Å	Õ	å	õ
&	6	F	V	f	v		¶	Æ	Ö	æ	ö
'	7	G	W	g	w		·	Ç	×	ç	÷
(	8	H	X	h	x		¸	È	Ø	è	ø
)	9	I	Y	i	y		¹	É	Ù	é	ù
*	:	J	Z	j	z		º	Ê	Ú	ê	ú
+	;	K	[	k	{		»	Ë	Û	ë	û
,	<	L	\	l			¼	Ì	Ü	ì	ü
-	=	M	]	m	}		½	Í	Ý	í	ý
.	>	N	^	n	~		¾	Î	Þ	î	þ
/	?	O	_	o	DEL		-	Ï	ß	ï	ÿ

SP: *Space* (Leerschritt, Zwischenraum) DEL: *Delete* (Löschen)

### Abb. 1.5-2: Genormte Zeichensätze

ECMA =  
*European Computer  
Manufacturers  
Association*, Genf

ANSI =  
*American National  
Standards Institute*,  
New York

ISO =  
*International  
Organization for  
Standardization*,  
Genf

DOS =  
*Disk Operating System*,  
ein Betriebssystem

Quelle: Zimmer 98, S. 291 ff./

*Interchange*) (Abb. 1.5-2) umfasst – grob gesagt – alle Zeichen, die sich auf einer englischen PC-Tastatur befinden, d.h., es gibt keine länderspezifischen Sonderzeichen wie Umlaute oder ß.



## LE 2 1.5 Dienste im Internet

Um Sonderzeichen innerhalb eines ASCII-Textes zu verwenden, wird das betreffende Sonderzeichen – in der Regel automatisch – in Klammern gesetzt und dort mit einem Zahlenwert codiert. Damit ist aber noch nicht das Problem gelöst, wie man Bilder oder Töne verschickt.

Zur Lösung dieses Problems wurden Anlagen bzw. Anhängsel (*attachments*) zur Nachricht ermöglicht. Sie erlauben es, an eine E-Mail jede andere Computerdatei anzuhängen. Damit können dann z.B. Word-Dokumente, Bilder und Ton-Dateien verschickt werden.

Anlagen werden binär im MIME-Standard (*multipurpose internet mail extensions*) übertragen. Die zu übertragenden Informationen werden beim Versand automatisch in ASCII-Zeichen codiert und beim Empfang wieder automatisch decodiert, vorausgesetzt beim Sender und Empfänger ist MIME installiert.

Antwort (*reply*) Will man eine erhaltene E-Mail beantworten, dann wird dies durch die Antwort-Funktion (*reply*) erleichtert. Sie bewirkt, dass automatisch eine E-Mail an den Absender erzeugt wird, bei der der Betreff beibehalten und – je nach Programm – besonders gekennzeichnet wird, z.B. durch ein vorangestelltes Re. Der Ursprungstext wird ebenfalls besonders gekennzeichnet. Antworten können einfach in den Ursprungstext eingefügt werden.

Weiterversenden (*forward*) Durch Weiterversenden (*forward*) kann eine E-Mail einem Dritten zugeleitet werden.

Netiquette  
Etikette =  
Benehmen, Brauch Im Internet gibt es eine Reihe von Verhaltensregeln, die man beachten sollte. Diese informellen Benimmregeln werden *Netiquette* genannt, eine Zusammensetzung aus Netzwerk und Etikette. Für den E-Mail-Dienst sollten folgende Regeln eingehalten werden:

- Keine irreführenden Betreff-Zeilen eintragen.
- Durchschläge (*cc*) nur wenn unbedingt nötig und nur an einen kleinen beschränkten Empfängerkreis. Nichts ist lästiger, als Unmengen unerwünschter Kopien zu sichten und zu löschen.
- Kurzfassen!

Mailing-Listen *Mailing-Listen* erweitern die elektronische Post. Eine E-Mail, die an eine *Mailing-Liste* geschickt wird, wird vervielfältigt und an alle Teilnehmer dieser Liste weitergeschickt. Diese können dann entweder an die Liste oder auch privat antworten. Es entsteht entweder ein Diskussionsforum oder nur ein Mechanismus zum Verteilen von E-Mails. Eine *Mailing-Liste* muss abonniert werden (*subscribe*), damit man an ihr teilnehmen kann. Alle über diese Liste verteilten Nachrichten landen im eigenen Postfach. Um eine *Mailing-Liste* zu verlassen, muss man das Abonnement rückgängig machen (*unsubscribe*).

Charakteristika Charakteristisch für den E-Mail-Dienst im Internet ist, dass das Zielcomputersystem die Post ohne Authentifizierung annimmt und für den Empfänger in einem besonderen Bereich bereithält. Der E-Mail-Dienst erlaubt den schnellen, asynchronen und informellen Austausch von Nachrichten. Der geographische Ort des Empfängers spielt keine Rolle mehr.

### 1.5.2 Nachrichtengruppen im Internet (*newsgroups*)

**Newsgrups** bzw. Nachrichtengruppen ermöglichen es, Computer-Konferenzen abzuhalten. Das bekannteste Computer-Konferenzsystem stellt das weltumspannende *Usenet* im Internet dar.

Bei einer Computer-Konferenz wird über elektronische »schwarze Bretter« asynchron kommuniziert. Zu verschiedenen Themengebieten gibt es Diskussionsforen. Jeder Teilnehmer kann auf die bisher eingebrachten Beiträge eines Forums zugreifen und bei Bedarf eigene hinzufügen.

asynchrone  
Kommunikation

Ähnlich wie bei *Mailing*-Listen basiert das *Usenet* technisch gesehen auf den Möglichkeiten der E-Mail. Jede einzelne Nachrichtengruppe besitzt eine eindeutige Adresse. Mitteilungen an eine Nachrichtengruppe werden an diese Adresse geschickt und dann wiederum allen Teilnehmern dieser Gruppe verfügbar gemacht. Jeder Teilnehmer kann einen Kommentar zu einer Mitteilung eines anderen Mitglieds an die Gruppe schicken, der an die Originalnachricht angehängt wird.

Ein weiterer Teilnehmer oder der Autor der ersten Nachricht können wiederum Stellung hierzu nehmen. Diese Nachricht wird dann ebenfalls an die bestehenden Mitteilungen angehängt usw. Einen solchen fortlaufenden Dialog bezeichnet man im *Usenet* als Faden (*thread*). Um an einer *Usenet*-Gruppe teilzunehmen, muss man sie abonnieren (*subscribe*).

Die Mitteilungen des *Usenet* werden mithilfe des *Network News Transport Protocols* (NNTP) im Internet befördert. Computersysteme, über die der Nachrichtenverkehr abgewickelt wird, bezeichnet man als NNTP-Server. Welche Nachrichtengruppen auf einem bestimmten NNTP-Server angeboten werden, entscheidet jeder Server-Betreiber für sich. Aus Kapazitätsgründen wird in der Regel nur eine begrenzte Anzahl angeboten. Es gibt jedoch viele öffentliche NNTP-Server, die über das Internet von jedem benutzt werden können.

NNTP

<http://newsserver.surfinfo.com>

Im *Usenet* gibt es über 50.000 verschiedene Nachrichtengruppen. Durch eine hierarchische Namensgebung werden die Gruppen gegliedert. Es gibt folgende sechs klassische Kategorien:

Kategorien der  
*Newsgrups*

- *comp* Themenbereiche Computersysteme und Informatik
- *news* Themen, die sich mit dem *Usenet* selbst befassen. Für Neueinsteiger sind folgende Gruppen nützlich:  
news.newusers.questions und news.announce.newusers.
- *sci* Themenbereiche Forschung (*science*) und Anwendung
- *soc* Gesellschaftliche und politische Diskussionen.
- *talk* Ort für ideologische, religiöse und sonstige kontroverse Debatten.
- *misc* Vermischtes (*miscellaneous*), d.h. alles, was in die anderen Hierarchien nicht passt.

## LE 2 1.5 Dienste im Internet

Eine Nachrichtengruppe beginnt in der Regel mit einem dieser Kürzel. Abgetrennt durch Punkte werden die Teile des Gruppennamens von links nach rechts immer spezieller, z.B. `comp.lang.java.programmer`.

Neben den klassischen Hierarchien gibt es viele so genannte alternative Hierarchien, für die weniger komplizierte Regeln gelten, um ein neues Forum einzurichten:

- **alt** Gruppen, die die Regeln der klassischen Kategorien umgehen. Beispiele für Gruppen sind:
  - `alt.binaries` Enthalten codierte Bilder, Audio-Dateien und Ähnliches
  - `alt.newusers.questions` Themen, die neue Mitglieder betreffen
  - `alt.test` Erlaubt es, Testmitteilungen zu verschickenIn dieser Hierarchie kann der Benutzer selbst neue Gruppen anlegen und löschen. In den klassischen Kategorien entscheidet dies ein Gremium.
- **de** Deutschsprachige Nachrichtengruppen. Unterhierarchien orientieren sich zum Teil an den klassischen Kategorien, z.B. `de.comp.security` oder `de.alt.drogen`
- **lokale** An jeder Universität existieren örtliche Nachrichtengruppen, die dann meist nur auf dem lokalen NNTP-Server vorhanden sind.



*Newsreader* Um an den Nachrichtengruppen im *Usenet* teilzunehmen, benötigt man als Software einen *Newsreader*. Die Benutzungsoberfläche eines *Newsreaders* gliedert sich in der Regel in die Bereiche (Abb. 1.5-3)

- Gruppenbereich,
- Nachrichtenbereich und
- Editor.

Im Gruppenbereich werden die abonnierten Nachrichtengruppen verwaltet. Alle auf dem jeweiligen NNTP-Server erhältlichen Gruppen können eingesehen und zum Abonnieren markiert werden.

Im Nachrichtenbereich werden die in den einzelnen abonnierten Gruppen vorhandenen Mitteilungen mit Absenderangabe, Datum und Uhrzeit des Versands sowie den Betreff-Angaben angezeigt.

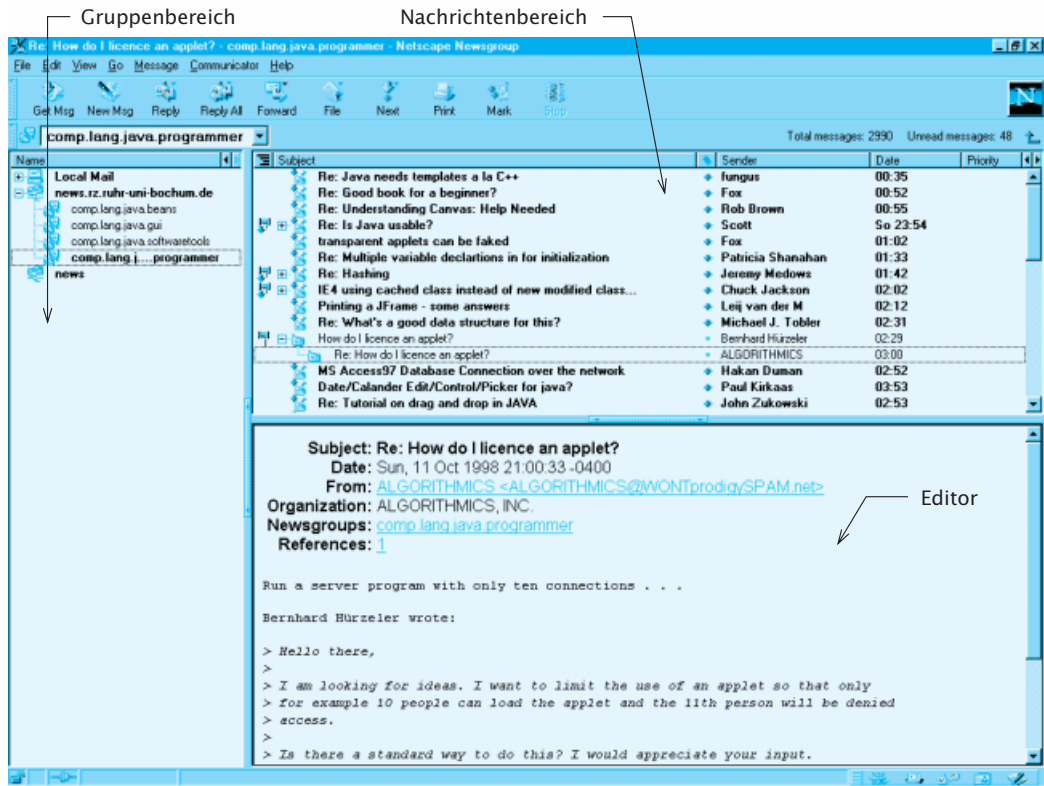
Um eine Nachricht zu lesen, wird sie in den Editor geladen. Am Ende der Mitteilung kann ein Kommentar angefügt werden (*follow up*), der zusammen mit der Mitteilung wieder an die Gruppe geschickt wird. Auf diese Weise entsteht ein »Faden« (*thread*).

Die meisten Browser besitzen heute einen integrierten *Newsreader*, z.B. der Internet Explorer.

Alternative zu *Newsreadern* Will man keinen *Newsreader* verwenden, dann kann man auch über Webportale gehen, die *Newsgroups* direkt im Web-Browser anzeigen, z.B. <http://groups.google.com>.

*Netiquette* Bei der Teilnahme an einer Nachrichtengruppe sollten folgende Benimmregeln befolgt werden:

## 1.5.2 Nachrichtengruppen im Internet (newsgroups) LE 2



- Aussagekräftige Betreff-Angaben wählen, damit die Mitglieder der Gruppe die Nachrichten schnell auf deren Inhalt hin durchsehen können.
- Kurzfassen!
- Diskussionen zuerst eine Weile mitverfolgen, um ein Gefühl für die Atmosphäre und die Themen der Gruppe zu erhalten. Erst dann mitreden (Fachbegriff: *posten*).
- Zuerst die Liste mit häufig gestellten Fragen durchlesen, um nicht dieselben Fragen zu stellen, wie viele andere Einsteiger vorher. Diese FAQs (*frequently asked questions*) werden regelmäßig in den entsprechenden Gruppen und/oder in news.answers veröffentlicht.
- Mitteilungen nur an die hierfür geeignete Nachrichtengruppe schicken.
- Alle Antworten auf Nachrichten (*follow ups*) lesen, um Wiederholungen zu vermeiden.
- In multikulturellen Gruppen mit Humor und Ironie vorsichtig sein.
- Niemals im Affekt schreiben, sondern jede Antwort vor dem Abschicken komplett durchlesen.
- Vollständige Sätze verwenden.

Abb. 1.5-3:  
Beispiel eines  
Newsreaders

FAQs

## LE 2 1.5 Dienste im Internet

- Vorsicht mit Abkürzungen (Akronymen), die der Diskussionsgruppe vielleicht nicht bekannt sind.
- Groß- und Kleinschreibung verwenden. WER NUR IN GROSSBUCHSTABEN SCHREIBT, SCHREIT!
- Gegenüber Rechtschreibschwächen tolerant sein. Es können auch Leitungsstörungen sein.
- Gegenüber Problemen in einer fremden Sprache tolerant sein.
- Aus der Diskussion sparsam zitieren (Fachbegriff: *quoten*).
- Immer daran denken: Etliche Tausende lesen mit.

### 1.5.3 Plaudern im Internet (*chat*)

*chat* Plaudern (***chat***) im Internet bedeutet, sich in nahezu Echtzeit mit anderen Benutzern zu unterhalten – im Gegensatz zur elektronischen Post und zu Nachrichtengruppen, die asynchron ablaufen.

IRC Der *Internet Relay Chat* (IRC) ist ein weltumfassendes, textorientiertes Kommunikationswerkzeug. Im IRC gibt es unterschiedliche Kanäle (*channels*), die sich um jeweils ein Thema drehen. Fast jede Universität hat ihren eigenen Kanal, der oft wie die Stadt heißt.

Um am IRC teilzunehmen, muss man sich bei einem IRC-Server anmelden.

Befehle Mit folgenden Befehlen kann man sich am Plaudern beteiligen:

- `/list` Anzeige aller Kanäle.
- `/list #b*` Anzeige aller Kanäle, die mit dem Buchstaben `b` beginnen.  
# kennzeichnet einen Kanalnamen.  
\* ist ein *wildcard*, das für eine beliebige Zeichenfolge steht.
- `/list *.ca` Anzeige sämtlicher Server, die mit der Endung `«.ca«` enden, wobei `ca` für Kanada steht.
- `/join #bochum` Anmelden bei einem Kanal. Es wird dann angezeigt, wer im Moment im Kanal angemeldet ist.
- `/join #neuerkanal` Einrichten eines neuen Kanals (bisher nicht verwendeter Name).

Ein Beispiel für eine Plauderei zeigt Abb. 1.5-4.

- `/me` Der eigene Name wird im *chat* vorangestellt, z.B. `/me geht es gut!` ergibt `* Helmut geht es gut!`
- `Name: Nachricht` Durch vorangestellten Namen angeben, wem die Nachricht gilt.
- `/part #bochum` Abmelden aus dem Kanal.

Webchat Eine Alternative zum IRC ist der Webchat. Er ist einfacher zu bedienen, denn er benötigt keine speziellen Programme und keine speziel-

*wildcard* = »wilde  
(d.h. beliebig  
verwendbare)  
Spielkarte«

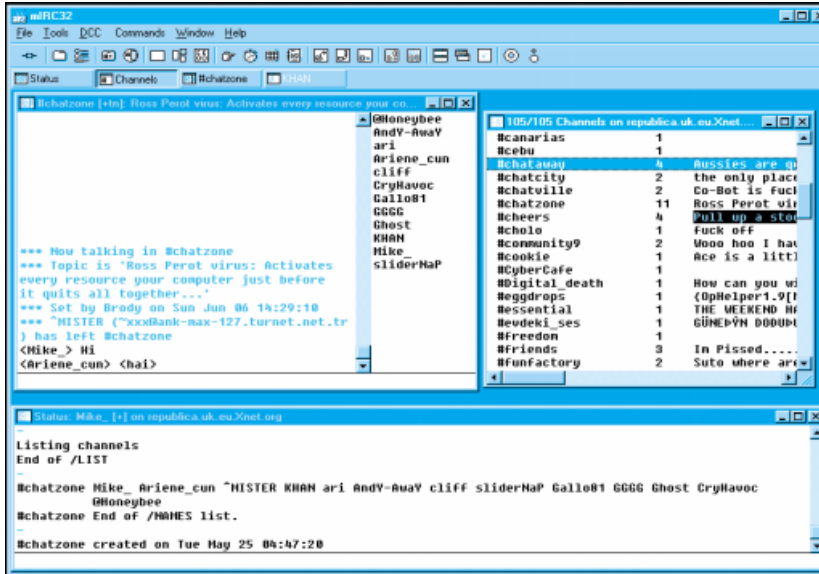


Abb. 1.5-4:  
Beispiel für ein  
IRC

len Kenntnisse. Will man an einem Webchat teilnehmen, dann muss man nur auf die Webseite eines entsprechenden Anbieters – z.B. [www.chatzentrale.de](http://www.chatzentrale.de) – gehen und kann mitmachen. Voraussetzung ist ein Java-fähiger Web-Browser. Im Browser müssen Java-Applets zugelassen sein.

### 1.5.4 Dateien übertragen im Internet (*ftp*)

Um Dateien zwischen Computersystemen im Internet zu übertragen, wurde das *file transfer protocol* (**ftp**) entwickelt. Damit ist es möglich, von *ftp*-Servern Informationen und Software auf das eigene Computersystem zu laden. Heute wird dieser Dienst auch intensiv dazu genutzt, aktualisierte Softwareversionen (*updates*) über das Netz von der Herstellerfirma zu beziehen.

Da *ftp*-Server oft auch Dateien bereitstellen, die nur für einen bestimmten Personenkreis zugänglich sein sollen, muss man sich aus Sicherheitsgründen auf dem gewünschten *ftp*-Server »ausweisen«.

In der Regel wird der Zugang zu Computersystemen durch die Vergabe einer so genannten *Log-in-ID* (Identifikation), die mit einem geheimen Passwort gekoppelt ist, geschützt. Nach dem Verbindungsaufbau werden diese beiden Angaben abgefragt. Diesen Vorgang bezeichnet man als Einloggen.

Log-in-ID  
Passwort

Einloggen

Um in den allgemein zugänglichen Bereich eines *ftp*-Servers zu gelangen, wird als *Log-in-ID* die Angabe »*anonymous*« oder »*ftp*« und als Passwort »*guest*« oder die eigene E-Mail-Adresse angegeben.

Damit die Übertragungszeit für Dateien minimiert wird, sind die Dateien in der Regel komprimiert abgespeichert. Man spricht von

## LE 2 1.5 Dienste im Internet

gepackten Dateien. Gepackte Dateien müssen nach der Übertragung entpackt werden.

packen, entpacken

Es gibt selbstentpackende Dateien, die ein ausführbares Programm enthalten, um die Datei zu dekomprimieren. Nicht selbstentpackende Dateien müssen durch das Programm entpackt werden, mit dem sie auch gepackt wurden.

Archie,  
Gopher

Um Dateien auf öffentlich zugänglichen *ftp*-Servern zu finden, kann man die Suchmaschinen Archie und Gopher einsetzen (siehe Abschnitt 1.6.2). Beide indizieren Dateien auf öffentlich zugänglichen *ftp*-Servern. Dadurch wird die Suche nach bestimmten Dateien und deren Lokalisierung möglich.

### 1.5.5 Computersysteme fernbedienen über das Internet (*Telnet*)

Der Internet-Client **Telnet** wurde entwickelt, um entfernte Computersysteme ferngesteuert über das Internet nutzen zu können. Dabei arbeitet das eigene Computersystem wie ein an ein zentrales Computersystem angeschlossenes *Terminal*.

*Terminal*

Unter einem *Terminal* versteht man einen Bildschirm und eine Tastatur, die mit einem entfernten Computersystem verbunden sind. Ein *Terminal* besitzt keinen eigenen Prozessor und damit auch keine Möglichkeit, Programme auszuführen. *Telnet* »emuliert« auf dem eigenen Computersystem ein zeichenorientiertes *Terminal*-Fenster.

Möglichkeiten

Mithilfe von *Telnet* ist es daher möglich, Änderungen auf einem entfernten Computersystem direkt vorzunehmen. Ohne *Telnet* müsste man die zu ändernde Datei erst auf den eigenen Computer laden, dort ändern und dann wieder auf den Zielcomputer übertragen.

Geben Sie in eine  
Suchmaschine ein:  
Tutorial Telnet



**Tim Berners-Lee**

Wegbereiter des Web am Kernforschungszentrum CERN in Genf (1990); heute: Direktor des WWW Consortiums und *principal research scientist* am MIT; Ausbildung: BA in Physik, *Oxford University*.

Außerdem können mit *Telnet* Programme direkt auf einem fremden Computersystem gestartet werden. Dies ist sehr nützlich, wenn man sich von einem fremden Computersystem auf ein Computersystem mit seinem eigenen E-Mail-Zugang einloggen will, um elektronische Post zu lesen.



Wer *Telnet* benutzen möchte, sollte sich zunächst Grundlagenwissen aneignen. So benötigt der *Telnet*-Benutzer die Kenntnis spezieller Unix-Kommandos.

## 1.6 Das World Wide Web (WWW)

Das **World Wide Web** (weltweites Netz), kurz **Web**, **W3** oder **WWW** genannt, ist ein

- globales,
- interaktives,
- dynamisches,
- plattformübergreifendes,

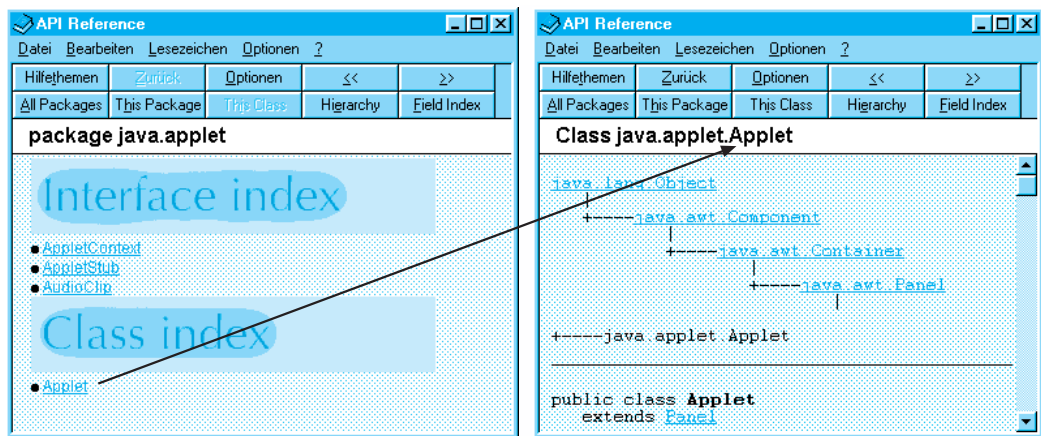


- verteiltes
- Hypermedia-Informationssystem, das auf dem Internet läuft.

Diese Definition wird im Folgenden näher erläutert.

Im Gegensatz zu den anderen Internetdiensten basiert das Web auf der Hypertext-Technik. Ein **Hypertext** ist ein Dokument, in dem wiederum auf ein oder mehrere andere Dokumente verwiesen wird. Die Verweise auf andere Dokumente bezeichnet man als **Hyperlinks** – kurz **Links** genannt. Ein solcher Verweis wird in der Regel durch ein farbig unterstrichenes Wort gekennzeichnet. Klickt man mit der Maus auf ein solches Wort, dann wird zu dem Dokument verzweigt, auf das der Verweis zeigt, d.h., dieses Dokument wird dann auf dem Bildschirm angezeigt. Durch diese Technik können Dokumente beliebig vernetzt werden. Hypertext-Dokumente müssen nicht mehr sequenziell gelesen werden. An jeder Verweisstelle kann zu anderen Dokumenten verzweigt werden, wenn man dies wünscht.

Das Hilfesystem, das das Betriebssystem Windows standardmäßig **Beispiel** anbietet, benutzt ebenfalls die Hypertext-Technik. Abb. 1.6-1 zeigt ein entsprechendes Beispiel.



Im Web kann nicht nur auf textorientierte Dokumente verwiesen werden, sondern auch auf Grafiken, Bilder, Videos, Ton usw. Daher spricht man von einem Hypermedia-System.

Während Hypertext-Hilfesysteme sich in der Regel vollständig auf einem Computersystem befinden, d.h., alle Hyperlinks verweisen auf Dokumente, die sich auf demselben Computersystem befinden, verweisen die Hyperlinks im Web oft auf Dokumente, die auf anderen Computersystemen im Internet gespeichert sind. Ein Mausklick auf einen Hyperlink im Web kann also dazu führen, dass ein Dokument von einem völlig anderen Computersystem – u.U. von einem anderen Kontinent – auf den eigenen Computer geladen wird.

Abb. 1.6-1: Beispiel für ein Hypertext-Hilfesystem

verteilt

## LE 2 1.6 Das World Wide Web (WWW)

Daher handelt es sich beim Web um ein verteiltes Informationssystem.

global, plattform-  
übergreifend

Da die Informationen weltweit verteilt und auf Computersystemen unterschiedlicher Hersteller mit unterschiedlichen Betriebssystemen liegen können, ist das Web global und plattformübergreifend.

dynamisch

Das Web ist dynamisch, da der Ersteller eines Web-Dokuments dieses jederzeit ändern kann, einschließlich der Verweise im Dokument. Umgekehrt gibt es aber keine Kontrolle über die Dokumente, die auf dieses Dokument verweisen. Wird ein Verweis nicht gefunden, dann wird eine Fehlermeldung ausgegeben.

interaktiv

Ein Web-Dokument kann Formulare enthalten, in die der Benutzer Informationen eintragen kann. Außerdem kann es Interaktionselemente wie Druckknöpfe oder Schalter enthalten. Daher ist das Web interaktiv.

### 1.6.1 Web-Browser

Adressierung

Die Adressierung im Web erfolgt nach denselben Prinzipien wie im Internet. Es kann alternativ die IP- oder die DNS-Adresse des gewünschten Web-Servers verwendet werden.

Für die Übertragung von Web-Dokumenten wird im Web das *hypertext transfer protocol* (http) verwendet.

Das Web erlaubt aber nicht nur den Zugriff auf Web-Dokumente, sondern vereint alle anderen Internetdienste unter einer einheitlichen Benutzungsoberfläche. Vom Web aus kann man also alle Möglichkeiten des Internet nutzen, ohne sich mit der speziellen Bedienung verschiedener Zugriffsprogramme vertraut machen zu müssen.

to browse =  
durchsehen,  
herumstöbern

Die Software, mit der man die Dienste des Web in Anspruch nehmen kann, bezeichnet man als **Web-Browser**.

Da Web-Browser nicht nur auf Web-Dokumente zugreifen können, sondern auch mit allen anderen Servern im Internet kommunizieren können, ist es notwendig, bei der Adressangabe mitzuteilen, welches Protokoll für die Übertragung zum gewünschten Computersystem verwendet werden soll.

Die Internet-Adresse, wie sie für einen Web-Browser verwendet wird, ist daher folgendermaßen aufgebaut:

protokoll://dns/verzeichnis /.../dateiname

Beispiel

<http://www.swt.ruhr-uni-bochum.de/multimedia/index.html>

In diesem Beispiel gibt **http** an, dass das *hypertext transfer protocol* verwendet werden soll. Doppelpunkt und zwei Schrägstriche (://) geben dem Browser an, dass jetzt die eigentliche Computeradresse folgt. **www** gibt an, dass es sich um den Web-Server handelt. Die Angabe **www** ist eine Konvention, von der aber oft abgewichen wird.

Optional kann nach der Computeradresse noch der komplette Pfad bis hin zu einem bestimmten Dokument angegeben werden.

Im Beispiel wird auf den Web-Server des Lehrstuhls für Software-Technik zugegriffen und dort auf die Datei `index.html` in dem Verzeichnis `multimedia`.

Will man über einen Browser auf einen ftp-Server zugreifen, dann lautet die allgemeine Adressenangabe:

`ftp://ftp.server`, z.B. `ftp://ftp.swt.ruhr-uni-bochum.de`

Die im Web verwendete, standardisierte Darstellung von Internet-adressen bezeichnet man als *uniform resource locator* (**URL**).

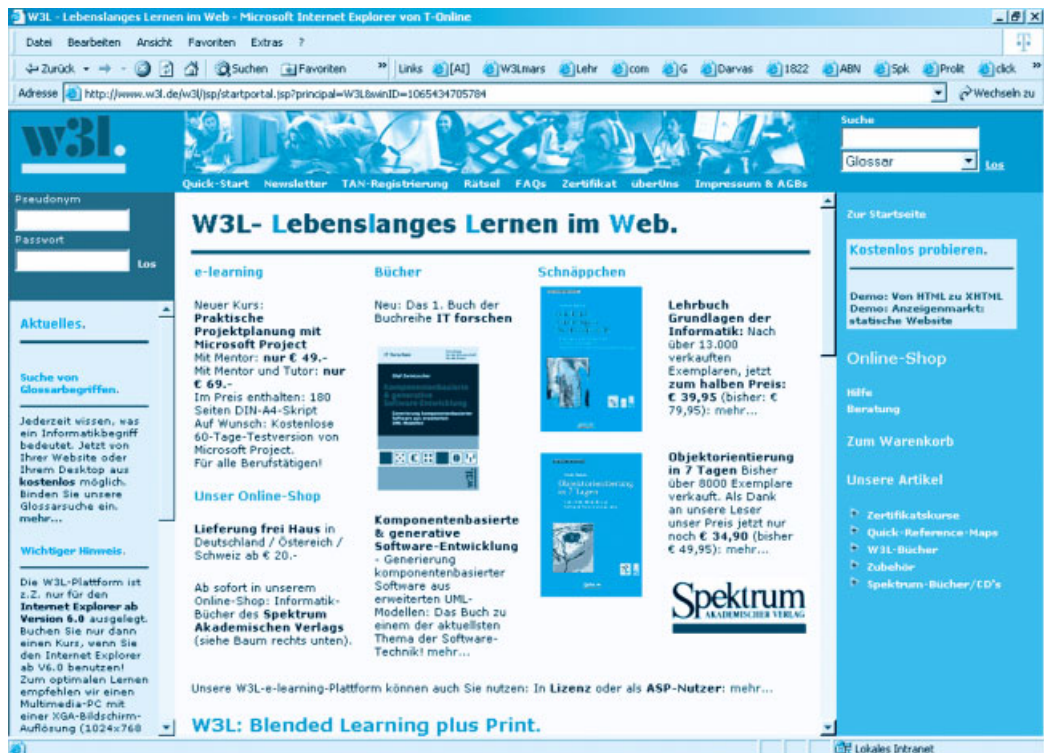
URL

Nach dem Start eines Web-Browsers erscheint auf dem Bildschirm ein Fenster mit Menüleiste, Druckknöpfen und einem Feld Adresse. Als Voreinstellung befindet sich in diesem Feld in der Regel bereits eine URL. Ist das Computersystem mit dem Internet verbunden, dann wird automatisch die erste Seite des angegebenen Web-Dokuments in das Computersystem geladen und angezeigt (Abb. 1.6-2).

Mehrere Web-Dokumente – auch **Webseiten** genannt – die untereinander durch Links sinnvoll miteinander verbunden sind, bilden eine so genannte **Website** – oft auch Web-Anwendung genannt (Abb. 1.6-2).

Die erste Seite einer Website bezeichnet man als **Start-** oder **Leitseite (Homepage)**. Durch Anklicken von Hyperlinks auf der Startseite gelangt man zu anderen Webseiten, die dann über das Netz

Abb. 1.6-2:  
Beispiel einer Web-  
seite, angezeigt in  
einem Web-Browser



## LE 2 1.6 Das World Wide Web (WWW)

in das Computersystem geladen werden. Dabei kann der Verweis auf einen völlig anderen Server zeigen. Dies merkt man unter Umständen überhaupt nicht.

### 1.6.2 Suchen und Finden im Internet

Im Internet befinden sich Unmengen an Informationen. Das Problem besteht darin, gezielt gewünschte Informationen zu suchen und zu finden.

Suchmaschine **Internet-Suchmaschinen (search engines)** durchsuchen automatisch alle verfügbaren Dateien (Abb. 1.6-3 und 1.6-4) und erstellen ein Verzeichnis der darin vorkommenden Wörter, allerdings ohne Füllwörter wie »und«, »der« usw. Diese Wörter werden alphabetisch geordnet als Index gespeichert, mit einem Hinweis, in welchen Dateien sie gefunden wurden.

Eine Anfrage mit Schlüsselwörtern liefert dann als Ergebnis alle Web-Adressen zurück, die hinter dem Schlüsselwort auftauchen.

allgemeine Suchstrategien Es gibt folgende allgemeine Suchstrategien:

- Festlegung von Schlüsselwörtern, die in einem engen Zusammenhang mit dem interessierenden Thema stehen.

Beispiel Es sollen Gestaltungshinweise für Webseiten gefunden werden. Schlüsselwörter sind: Gestaltung, Webseite, WWW-Seite, Homepage, design.

- Um die Menge der gefundenen Adressen zu reduzieren, kann oft auch nur in Zusammenfassungen (*abstracts*) gesucht werden, da dort oft nur für das Thema relevante Wörter verwendet werden. Webseiten selbst verfügen über keine Zusammenfassung, jedoch werden die Titelzeile (<title>, siehe nächstes Kapitel) und unsichtbare Schlüsselwörter (*keywords*) ausgewertet.
- Eine Sucheinschränkung ist auch durch die Angabe des Erstellungsdatums oder des Erstellungszeitraums möglich.
- Durch die Reduktion auf den Wortstamm (*trunk*), kann nach beliebigen Wortendungen, aber auch Wortanfängen gesucht werden. Mithilfe von Sonderzeichen, so genannten »wildcards« werden Platzhalter für beliebige Buchstaben geschaffen. Im Allgemeinen wird das ? für genau einen oder keinen Buchstaben und der \* für eine beliebige Anzahl von Buchstaben verwendet.

Beispiel Eine Suchanfrage mit »Gestalt\*« wird Texte mit Worten wie Gestaltung, Gestalten usw. liefern.

- Mehrere Suchbegriffe können mit den booleschen Operatoren **and**, **or** und **not** verknüpft werden.

Beispiel Gestaltung **and** (Webseite **or** WWW-Seite **or** Homepage)

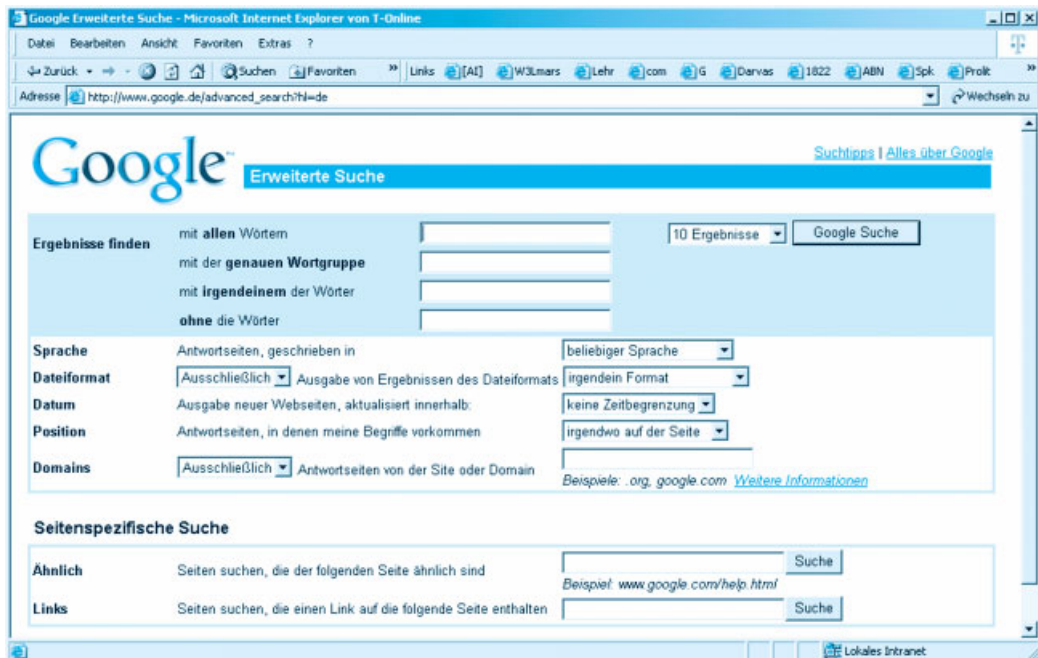


In Abhängigkeit von der Art der gesuchten Information gibt es unterschiedliche Möglichkeiten:

- Anregungen und Ideen: *Newsgroups* oder *Mailing*-Listen.  
Übersicht über viele öffentlich zugängliche *Mailing*-Listen:  
<http://www.mun.ca./cc/tsg/lists/index.html>  
Einstieg bei *Newsgroups*: [de.newusers.questions](http://de.newusers.questions)
- Referenzinformationen  
Literatursuche: Nordrhein-westfälischer Bibliotheksverbund:  
<http://www.hbz-nrw.de>  
Südwestdeutscher Bibliotheksverbund:  
<http://www.swbv.uni-Konstanz.de/CGI/cgi-bin/opacform.cgi>  
*Library of Congress* USA:  
<http://lcwww.loc.gov/homepage/index.html>
- Fakteninformationen: *ftp* oder Web  
Empfehlenswert ist der Einstieg über Institutionen oder Organisationen, die sich mit entsprechenden Themen befassen.  
Wissenschaftliche Informatikorganisationen:  
*ACM (Association for Computing Machinery)*  
<http://www.acm.org>  
*IEEE (Institute of Electrical and Electronics Engineers)*  
<http://www.computer.org>  
*GI (Gesellschaft für Informatik)*  
<http://www.gi-ev.de>

Außerdem kann über Suchmaschinen gesucht werden (Abb. 1.6-3).

Abb. 1-6-3:  
Beispiel für die  
»Homepage« einer  
Suchmaschine



**Abb. 1.6-4:  
Web-Suchdienste**

- Kataloge: Systematisch nach bestimmten Themengebieten angelegt. Klein (meist bis 100.000 Einträge), aber genau.
- Suchmaschinen (*search engines*, *web-robots*, *spiders*): Softwaresysteme, die sich entlang der *Hyperlinks* ziellos durch das Web »hangeln« und alle Seiten in eine Datenbank aufnehmen, die sie auf ihrem Weg finden.
- Nur Web.
- Auch *ftp*, *Newsgroups* usw.
- Metasuchsysteme: Schicken eine Anfrage gleichzeitig an mehrere Suchdienste.

**Beispiele**

- Yahoo! <http://www.yahoo.com>, <http://www.yahoo.de>  
Bester Suchdienst für denjenigen, der erst einmal wissen will, was es zu einem bestimmten Thema alles gibt.  
Sucht auf Wunsch nicht nur Webseiten, sondern auch *Newsgroups* und E-Mail-Adressen. Suche kann auf Rubriken beschränkt werden.
- Alta Vista: <http://www.altavista.com>  
Schnelle Suche, riesiger Datenbestand. Für alle, die genau wissen, was sie suchen.  
Die Abfrage  
+ link: <http://my.site.de/> -url: <http://my.site.de/>  
erlaubt es festzustellen, wie viele Hyperlinks von außerhalb auf die Webseite <http://my.site.de/> zeigen.
- Who Where? <http://www.whowhere.com>  
Speziell zum Auffinden von Personen, Organisationen, E-Mail-Adressen, Anschriften, Telefonnummern.
- Virtual Library (CERN): <http://www.vlib.org>  
Für wissenschaftlich und wirtschaftlich Interessierte.
- Lycos: <http://www.Lycos.com>, <http://www.Lycos.de>  
Umfangreich, aber oft überlastet.
- Webcrawler: <http://webcrawler.com>  
Übersichtliche Suche.
- Hotbot: <http://www.hotbot.com>  
Hervorragendes Abfrageformular, getrennte Suche nach Personen und Netzadressen.
- Excite: <http://www.excite.com>, <http://www.excite.de>  
Suche auch nach längeren Begriffen.
- Magellan: <http://www.mckinley.com>  
Durchsucht auch *ftp* und *Newsgroups*.
- Dejanews: <http://www.dejanews.com>  
Spezialisiert auf *Newsgroups*.
- Kolibri: <http://www.kolibri.de>  
Hohe Aktualität, sinnverwandte Begriffe möglich.
- Web.de: <http://vroom.web.de>
- Dino: <http://www.dino-online.de>  
Deutscher Katalog-Suchdienst.
- Fireball: <http://www.fireball.de>  
Deutscher Suchdienst, kennt deutsche Sonderzeichen und Umlaute.

**Metasuchsysteme**

- All-in-one: <http://www.albany.net>  
Erlaubt es, fast alle existierenden Suchmaschinen abzufragen, auch jene für E-Mail-Adressen.
- Metacrawler: <http://www.metacrawler.com>
- Dogpile: <http://www.dogpile.com>

**Übersicht über Suchmaschinen:**

- [http://www.ub2.l.u.se/tk/websearch\\_systemat.html](http://www.ub2.l.u.se/tk/websearch_systemat.html)
- <http://www.searchenginewatch.com>



Schwieriger ist die Suche nach Dateien auf *ftp*-Servern. Hier helfen Programme wie Archie, Gopher oder ftpsearch.

■ Personen und Institutionen, Web

Suche mithilfe einer Suchmaschine (Abb. 1.6-3 und Abb. 1.6-4). Zusätzliche Hilfe bei der Suche bieten spezialisierte Softwaresysteme, die *Web-Robots* und *Software-Agents* (Handelnde, Reisende) genannt werden.

Das Web wurde 1991 am europäischen Kernforschungszentrum CERN in Genf entwickelt. Ziel war es, den Zugriff auf die verschiedenen Internet-Dienste über eine einheitliche Benutzeroberfläche zu ermöglichen. Um eine große Kompatibilität zwischen verschiedenen Computerplattformen zu ermöglichen, wurde HTML entwickelt. Für diese Sprache ist das später dafür gegründete WWW Consortium (W3C) zuständig. Die aktuelle Version ist HTML 4.01.

## 1.7 Die Sprachen HTML und XHTML

Der Aufbau und die Struktur von Websites wird mit der Sprache HTML (*Hyper Text Markup Language*) beschrieben.

**HTML** ist eine Dokumentenauszeichnungssprache und bietet die Möglichkeit, inhaltliche Kategorien von Dokumenten durch HTML-Befehle zu kennzeichnen.

Inhaltliche Kategorien sind z.B.

- sechs verschiedene Überschriftsebenen (*headings*),
- Abschnitte (*paragraphs*),
- Aufzählungen (*bulleted lists*).

Jede inhaltliche Kategorie wird in entsprechende HTML-Formatierungsbefehle (*tags*) »eingebettet«. Beim Laden eines so gekennzeichneten Dokuments in einem Web-Browser formatiert der Browser das Dokument und stellt es entsprechend dar.

tags



Im Jahr 2000 wurde als Nachfolger von HTML die Sprache XHTML 1.0 (*eXtensible HyperText Markup Language*) als Web-Standard verabschiedet. XHTML enthält alle Befehle von HTML 4.01, besitzt aber eine strengere Syntax. Ziel von XHTML ist es, HTML abzulösen. Im Folgenden wird immer XHTML erklärt und verwendet. Eine ausführliche Einführung in HTML, XHTML und CSS findet man z.B. in /Balzert 03a/, einen Überblick über alle wichtigen Sprachelemente in /Balzert 03b/.

XHTML-Referenz  
<http://www.w3.org/TR/xhtml1/>

Tab. 1.7-1 zeigt die wichtigsten Formatierungsbefehle von XHTML. Abb. 1.7-1 stellt einen Web-Text mit Formatbefehlen und die formatierte Darstellung dieses Textes in einem Web-Browser gegenüber.

Wie die Abb. 1.7-1 zeigt, erfolgt die Darstellung des gewünschten Formats – im Unterschied zu einem Textverarbeitungssystem – nicht bei der Erstellung des Dokuments, sondern erst, wenn dieses in einen Web-Browser geladen wird. Beim Laden des XHTML-Dokuments interpretiert der Web-Browser die XHTML-Befehle und setzt sie in eine entsprechende Darstellung um. Verschiedene Browser stellen das XHTML-Dokument unterschiedlich dar, d.h. die Formatierungsbefehle müssen nicht von jedem Browser gleich umgesetzt werden.

Die Anzeige der für Hyperlinks definierten Textteile erfolgt in der hierfür im jeweiligen Browser festgelegten Darstellung, z.B. in der Farbe blau und unterstrichen.



**Tab. 1.7-1:**  
**Die wichtigsten**  
**XHTML-Befehle**

Jedes XHTML-Dokument setzt sich prinzipiell aus **vier Abschnitten** zusammen:

- der Deklaration des Dokumententyps,
  - dem Wurzelement: `<html>` Dokumentinhalt `</html>`
  - dem Kopfteil (*header*): `<head>` Kopfteil `</head>`
  - dem eigentlichen Dokument, Rumpf genannt (*body*): `<body>` Hauptteil `</body>`
- Jeder XHTML-Befehl ist in spitzen Klammern eingeschlossen. XHTML-Befehle sind symmetrisch aufgebaut, d.h., es gibt einen Anfangsbefehl und einen Abschlussbefehl. Der Abschlussbefehl wiederholt den Anfangsbefehl und enthält einen vorangestellten Schrägstrich. Durch Attribute können spezielle Eigenschaften für den Befehl festgelegt werden.

**Grundgerüst** eines vollständigen XHTML-Dokuments:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

```
<html>
<head>
<title> Titelzeile des Dokuments </title>
</head>
<body>
Eigentlicher Inhalt des Dokuments, der verschiedene
Formatierungsanweisungen enthalten kann
</body>
</html>
```

Die **Titelzeile** erscheint nicht im Dokument selbst, sondern – in Abhängigkeit vom jeweiligen Browser – beispielsweise in der Titelzeile des Web-Browsers. Außerdem wird die Titelzeile von Suchmaschinen ausgewertet.

#### Überschriften (*headings*)

```
<h1> Überschriftsebene 1 </h1>
bis
<h6> Überschriftsebene 6 </h6>
```

#### Abschnitte (*paragraphs*)

```
<p> Abschnitt </p>
```

Es wird eine Leerzeile eingefügt, Zeilenumbrüche haben keine Wirkung.

#### Hervorhebungen (*formatting*)

```
<b> Fetter Text (bold) </b>
<i> Kursiver Text (italic) </i>
<u> Unterstrichener Text
      (underline) </u>
<hr/> Horizontale Linie (horizontal rule)
      (Achtung: Implizierter Abschluss-
      befehl durch /)
<br/> Zeilenumbruch (break)
      (Achtung: Implizierter Abschluss-
      befehl durch /)
```

Wird das html-Dokument mit einem Texteditor erstellt, dann ist die Datei als ASCII-Datei mit der Endung .html abzuspeichern!

```
<!--kommentar-->
```

#### Aufzählungen (*unordered lists*)

```
<ul>
<li> Listenelement (list item)</li>
<li> Listenelement </li>
...
</ul>
```

#### Numerische Aufzählungen (*ordered lists*)

```
<ol>
<li> Listenelement </li>
<li> Listenelement </li>
...
</ol>
```

#### Hyperlink einbinden (*anchor*)

```
<a href = "URL">
Zielangabe für den Benutzer
</a>
href="URL" ist ein Attribut mit dem Wert
"URL"
Beispiel für Zugriff auf lokale Datei in
Windows:
<a href = "file://localhost/
c:\html\
Beispiel.html">
Ein Beispiel
</a>
```

#### Bild einbinden (*image*)

```
<img align = "top" oder "middle"
oder "bottom" oder "left" oder
"right" src = "URL"/>
```

## 1.7 Die Sprachen HTML und XHTML LE 2

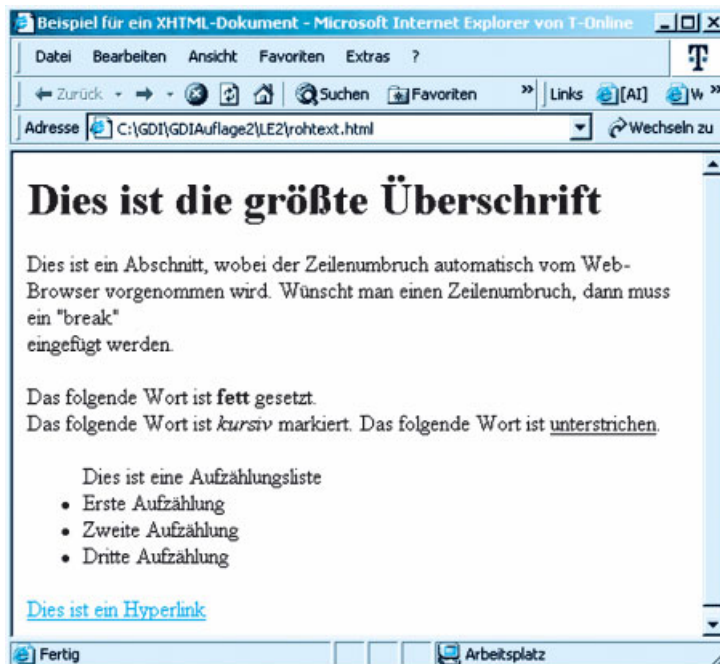
Jeder Web-Browser verfügt über einen Befehl (z.B. *Open Local File*), um »lokale« XHTML-Dokumente, die sich auf demselben Computersystem befinden, zu laden.



```
rohtext.html - Editor
Datei Bearbeiten Format ?

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html>
  <head>
    <title>Beispiel für ein XHTML-Dokument</title>
  </head>
  <body>
    <h1>Dies ist die größte Überschrift</h1>
    <p>Dies ist ein Abschnitt, wobei der
    Zeilenumbruch automatisch vom Web-Browser vorgenommen wird.
    Wünscht man einen Zeilenumbruch,
    dann muss ein "break" <br/>eingefügt werden.
    </p>
    <p>Das folgende Wort ist <b>fett</b> gesetzt.<br/>
    Das folgende Wort ist <i>kursiv</i> markiert.
    Das folgende Wort ist <u>unterstrichen</u>.
    </p>
    <ul>Dies ist eine Aufzählungsliste
    <li>Erste Aufzählung</li>
    <li>Zweite Aufzählung</li>
    <li>Dritte Aufzählung</li>
    </ul>
    <a href="Blume.jpeg">
    Dies ist ein Hyperlink</a>
  </body>
</html>
```

Abb. 1.7-1:  
HTML-Rohtext und  
seine Browser-  
Darstellung



## LE 2 1.7 Die Sprachen HTML und XHTML

### 1.7.1 Grafiken und Bilder einbinden

XHTML-Dokumente enthalten ausschließlich Text. Andere Medien wie Grafiken, Bilder, Audios und Videos können im Unterschied zu einem komfortablen Textverarbeitungssystem nicht direkt in das Dokument eingebunden werden.

An der gewünschten Stelle wird vielmehr ein Hyperlink auf die entsprechende Datei gesetzt. Der entsprechende XHTML-Befehl für das Einbinden von Grafiken und Bildern lautet:

```
<img align = "top" oder "middle" oder "bottom" oder "right" oder "left" src = "URL"/>
```

Hinweis:  
Das Attribut `align`  
gehört zu HTML

Hinter `align` (*alignment*) wird angegeben, ob der auf diesen Befehl folgende Text rechts am oberen Rand (*top*) der Grafik bzw. des Bildes ausgerichtet wird oder in der Mitte (*middle*) oder am unteren Rand (*bottom*) oder am rechten Rand (*right*) oder am linken Rand (*left*) (Abb. 1.7-2). Die `align`-Angabe bezieht sich nur auf die Textzeile, in der das Bild bzw. die Grafik aufgeführt ist.

Die meisten Browser unterstützen in der Grundausstattung nur die Grafik-/Bildformate `gif` und `jpg` (Abb. 1.7-3). Daher müssen andere Grafikformate vorher entsprechend konvertiert werden.

Steht ein `<img>`-Befehl innerhalb eines `<a>`-Befehls, dann wird nach Anklicken des Bildes zur Verweisstelle verzweigt, z.B.

```
<a href = "Startseite.html"><img src = "Pfeil.gif"/></a>
```

Damit ist es möglich, Piktogramme zum Navigieren innerhalb einer Website einzusetzen.

Auf Webseiten ist es üblich, von Bildern und Grafiken eine verkleinerte Form als Vorschau zu erstellen, damit der Benutzer schon eine Vorstellung von dem Bild bekommt, bevor – wegen der Übertragungszeit – das ganze Bild erscheint.

*thumbnails*

Diese briefmarkengroßen Miniaturen werden Daumennägel (*thumbnails*) genannt. Durch Anklicken wird dann das Vollbild angezeigt.

Beispiel 

```
<a href = "grossesBild.gif"><img src = "kleinesBild.gif"/></a>
```

Am Ende der Startseite sollte angegeben werden, wer für die Website verantwortlich oder zuständig ist. Diese Person bezeichnet man als »*webmaster*«. Außerdem sollte der Zustand des Dokuments aufgeführt werden (fertig gestellt, im Aufbau). Wird eine Website oft geändert, dann sollte das letzte Änderungsdatum angezeigt werden. Wünschenswert ist es, auf der Webseite einen Verweis zur E-Mail-Adresse des *webmasters* anzubringen, z.B. in folgender Form:

```
address <address>
Helmut Balzert <a href = "mailto:hb@swt.ruhr-uni-bochum.de" >
hb@swt.ruhr-uni-bochum.de </a>
</address>
```

## 1.7.1 Grafiken und Bilder einbinden LE 2

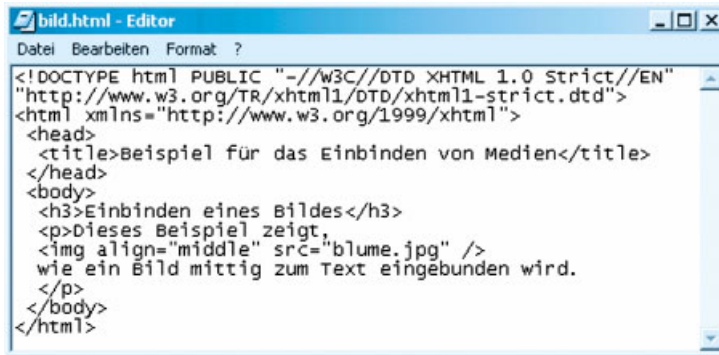
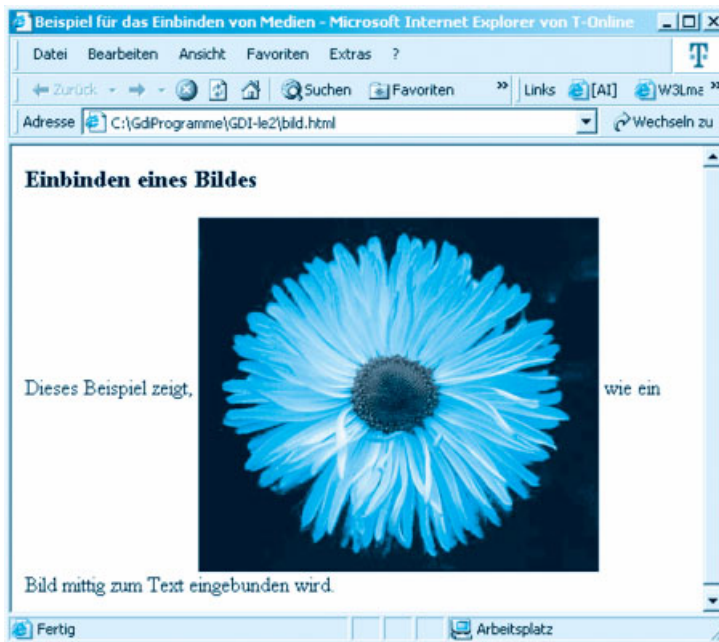


Abb. 1.7-2:  
Beispiel für das  
Einbinden von  
Grafiken und  
Bildern



Piktogramm-Sammlungen sind zu finden unter:  
[http://dir.yahoo.com/Arts/Design\\_Arts/Graphic\\_Design/Web\\_Page\\_Design\\_and\\_Layout/Graphics/](http://dir.yahoo.com/Arts/Design_Arts/Graphic_Design/Web_Page_Design_and_Layout/Graphics/)  
<http://www.sct.gu.edu.au/~anthony/icons/index.html>  
<http://www.iconarchive.com/>

Dieser Befehl ermöglicht es, direkt eine E-Mail an die betreffende Person zu senden, ohne die Adresse nochmals ins E-Mail-Programm eingeben zu müssen.

HTML unterstützt ursprünglich nur den unteren Teil des ISO-Latin1-Zeichensatzes (siehe Abb. 1.5-2, unten, linke Seite der Tabelle).

Damit der gesamte ISO-Latin1-Zeichensatz verwendet wird, ist folgender HTML-Befehl in den Kopf des HTML-Dokuments einzutragen:

```
<meta http-equiv = "Content-Type"
content = "text/html"; charset = "iso-8859-1">.
```

Die benutzergerechte Gestaltung einer Website gehört zu dem Gebiet der »Software-Ergonomie« (siehe z.B. /Balzert 04/).

Sonderzeichen

Übersicht über  
Sonderzeichen  
und Codes:  
<http://www.december.com/html/>

**Abb. 1.7-3:  
Grafikformate  
im Web**

Alles über gif:  
<http://members.aol.com/royalef/gifabout.htm>

<http://www.zampano.com/gifanim/index0.html>

Für Netzübertragungen wurden Grafikformate entwickelt, die Bilder und Grafiken so komprimieren, dass sie möglichst wenig Speicherplatz benötigen und daher kurze Übertragungszeiten erlauben.

**gif-Format** (*graphics interchange format*) (Dateiendung .gif)

- Für Grafiken, die wenige optische Informationen enthalten, z.B. Logos, Banner, Bilder mit Schrift.
- Ein gif-Bild besitzt eine Farbpalette, die bis zu 256 Farben (8 Bit pro Pixel) enthalten kann. Je kleiner die Farbtabelle, desto geringer ist auch der Platzbedarf.
- Die Einstellung auf eine begrenzte Farbtabelle darf erst am Ende erfolgen, da mit der Reduzierung der Farben Daten verloren gehen.
- Außer einer Reduzierung der Farbpalette verwendet gif zusätzlich eine LZW-Kompression. Sie arbeitet zeilenweise und bleibt daher bei horizontalen Farbverläufen wirkungslos.

**interlaced gif-Format**

- Die Zeilen eines Bildes werden nicht nacheinander dargestellt, sondern, ähnlich wie bei einer Fernsehröhre, versetzt.
- Ist in der Datei festgelegt, dass zuerst jede fünfte Zeile gezeigt wird, dann verringert sich die Ladezeit entsprechend. Jedoch ist das Bild dann zunächst unscharf zu sehen.

**jpeg-Format** (*joint photograph expert group*) (Dateiendung .jpg oder .jpeg)

- Für Fotografien. Für Strichzeichnungen oder Text in Bildern *nicht* geeignet.
- Kommt bei fotografischen Motiven der menschlichen Sichtweise entgegen.
- Der DCT-Kompressionsalgorithmus (*discrete cosinus transformation*) reduziert die Daten, ohne dass die Wahrnehmung wesentlich beeinträchtigt wird, obwohl auch der Inhalt des Bildes verändert wird.
- Ein Bild sollte **erst nach der Fertigstellung in das jpeg-Format** gewandelt werden, da Informationen verloren gehen. Jedes mal wenn ein jpeg-Bild bearbeitet wird, verschlechtert sich die Qualität.
- Kann ungefähr **16,7 Millionen Farben** enthalten.
- Die Farbanzahl in einem JPEG-Bild darf *nicht* reduziert werden, da sich die Bildgröße dann u.U. erhöht.

**Progressives jpeg-Format**

- Erlaubt mehrere »scans« in einem Bild, d.h., nach und nach baut sich das Bild als Mosaik aus quadratischen Feldern auf. Sie werden beim Laden immer kleiner und schärfer.
- Die Dateigröße nimmt nur minimal zu.

**png-Format** (*portable networks graphics*) (Dateiendung .png)

- Eignet sich für Animationen.
- Einzelbilder können in Abhängigkeit voneinander gespeichert werden.
- Kann Millionen von Farben enthalten.



Quelle: /Meissner 96/

**1.7.2 Erweiterungen von Web-Browsern (plug-ins)**

Web-Browser waren ursprünglich nur in der Lage, HTML-Dokumente in eine geeignete Darstellung zu überführen und anzuzeigen.

Helper  
Applications

Die älteste Erweiterungsmöglichkeit sind die so genannten *Helper Applications* (Hilfsanwendungen) oder *External Viewers* (externe Betrachter). Dies sind eigenständige Programme auf demselben Computersystem, die der Browser aufruft, wenn ein Dokument einem passenden MIME-Type entspricht. MIME ist die *multipurpose internet multimedia extension* und ermöglicht es, einer Datei eine Kategorie und einen Untertyp zuzuordnen.

## 1.7.2 Erweiterungen von Web-Browsern (*plug-ins*) LE 2



Abb. 1.7-4:  
Bilder als  
Verweisstellen



Hinweise zur Web-  
Gestaltung erhält  
man z.B. auf  
folgender Netz-  
adresse:  
[http://web.  
glover.com/](http://web.glover.com/)

Ein Bild hat den MIME-Typ Image und Untertypen wie z.B. gif und jpeg. Beispiel

Kommen die Daten von einem Web-Server, dann ist der MIME-Typ im HTTP-Vorspann angegeben, sonst gibt es eine Zuordnung zu einem Dateinamen-Suffix, z.B. .gif. Welches Programm solch ein Dokument darstellen soll, ist im Web-Browser festgelegt.

*External Viewers* stellen die Daten *nicht* innerhalb des Web-Browsers dar, sondern übernehmen selbst die Kontrolle über ihren weiteren Ablauf. Die Vorgehensweise ist relativ sicher. Es werden keine Programme auf das eigene Computersystem geladen, sondern nur Daten. Nachteilig ist, dass es keine standardisierte Rückkopplung zum aufrufenden Web-Browser gibt.

*External Viewers*

Um diese Nachteile zu vermeiden, wurden von der Firma Netscape so genannte *plug-in*-Schnittstellen (Einsteck-Schnittstellen) entwickelt. Es handelt sich bei den *plug-ins* um Anwendungen. Sie

*plug-ins*



## LE 2 1.7 Die Sprachen HTML und XHTML

können mit dem Netscape Navigator, dem Web-Browser der Firma Netscape, über definierte Schnittstellen kommunizieren und sind in eine Webseite integrierbar.

Der Einsatzzweck bestand ursprünglich darin, herstellereigene Dokumentformate statt mit einem externen *Viewer* direkt im Browser zu betrachten. Inzwischen gibt es aber auch komplette interaktive Anwendungen.

*Plug-ins* werten MIME-Typen aus. Sind mehrere dem gleichen Typ zugeordnet, dann gilt das zuletzt eingetragene *plug-in*. Dieses Eintragen, Registrierung genannt, erfolgt beim Start des Navigators. Welche *plug-ins* bekannt sind, zeigt die interne URL `about:plugins`. *Plug-ins* haben uneingeschränkten Zugriff auf den Web-Client.

Viele *plug-ins* für Netscape sind heute plattformübergreifend verfügbar.

Die bekanntesten *plug-ins* sind:

- *Shockwave/Flash*: Ermöglicht es, Multimedia-Anwendungen, die mit Software der Firma Macromedia erstellt wurden, abzuspielen.
- *pdf (portable document format)*: In diesem Format dargestellte Dokumente können angezeigt werden.

Mithilfe von *plug-ins* ist es auch möglich, Radio (Abb. 1.7-5) und Video bzw. Fernsehen über das Internet zu hören bzw. zu sehen. Auch ist es möglich, so genannte Kanäle (*channels*) zu abonnieren und Webseiten auf seine eigenen Bedürfnisse zu personalisieren.

Im Gegensatz zu *plug-ins* haben die meisten Web-Browser einen Interpreter für Java-Programme fest integriert.

Abb. 1.7-5:  
Beispiel für Radio  
über das Internet



### 1.7.3 Erstellen von XHTML-Seiten und Bereitstellen auf Web-Servern

Folgende Möglichkeiten gibt es, XHTML-Seiten zu erstellen:

- Einsatz eines Texteditors und Verwendung der XHTML-Befehle,
- Einsatz eines XHTML-Konverters bzw. Filters, der ein Dokument in ein XHTML-Dokument umwandelt,

Überblick über  
HTML-Werkzeuge:  
<http://www.w3.org/tools>





### 1.7.3 Erstellen von XHTML-Seiten und Bereitstellen auf Web-Servern LE 2

- Erweiterung von Standardanwendungen, z.B. Textverarbeitungssystemen, um XHTML-Zusätze (*add-ins*),
- XHTML-Editoren,
- Web-Browser mit integrierten XHTML-Editoren und
- vollständige Web-Entwicklungsumgebungen (Editor, Web-Server-Verwaltungswerkzeug und Web-Server-Software).

Sind die XHTML-Seiten erstellt, dann müssen sie auf einen Web-Server übertragen werden, damit sie im Internet allgemein zugänglich sind. Folgendes ist zu tun:

- 1 Alle zur Website gehörenden Dateien sind in ein Hauptverzeichnis zu legen. Unterverzeichnisse können für Grafiken usw. angelegt werden. ein Hauptverzeichnis
- 2 Alle Verweise in den Dateien sollten relativ zum Hauptverzeichnis gesetzt werden. Dadurch kann man das Hauptverzeichnis auf verschiedene Server legen, ohne dass die Verweise ungültig werden. Relative Verweise beginnen mit dem Datei- oder Unterverzeichnis-Namen, absolute Verweise beginnen mit einem Schrägstrich »/«.

<code>href = "datei.html"</code>	datei.html befindet sich im aktuellen Verzeichnis.	Beispiele
<code>href = "unterverzeichnis/datei.html"</code>	datei.html befindet sich im Verzeichnis unterverzeichnis und dieses im aktuellen Verzeichnis.	
<code>href = "../datei.html"</code>	datei.html befindet sich im Verzeichnis oberhalb des aktuellen Verzeichnisses – dasselbe Verzeichnis, in dem das aktuelle Verzeichnis ist.	

- 3 Die Start-Seite bzw. der oberste Index für jede Website sollte den Dateinamen index.html erhalten, da diese Datei von vielen Browsern als Voreinstellung geladen wird, wenn die URL keine explizite Startseitenangabe enthält. Dateinamen Startseite: index.html
- 4 Überprüfen, dass jede Datei durch den richtigen Dateityp gekennzeichnet ist, z.B. `Ton.au`.
- 5 Übertragen des gesamten Hauptverzeichnisses an die vorgegebene Stelle auf dem Web-Server. Die konkrete Übertragungsart hängt vom Server ab.

## 1.8 Fallstudie: Die Firma ProfiSoft im Internet

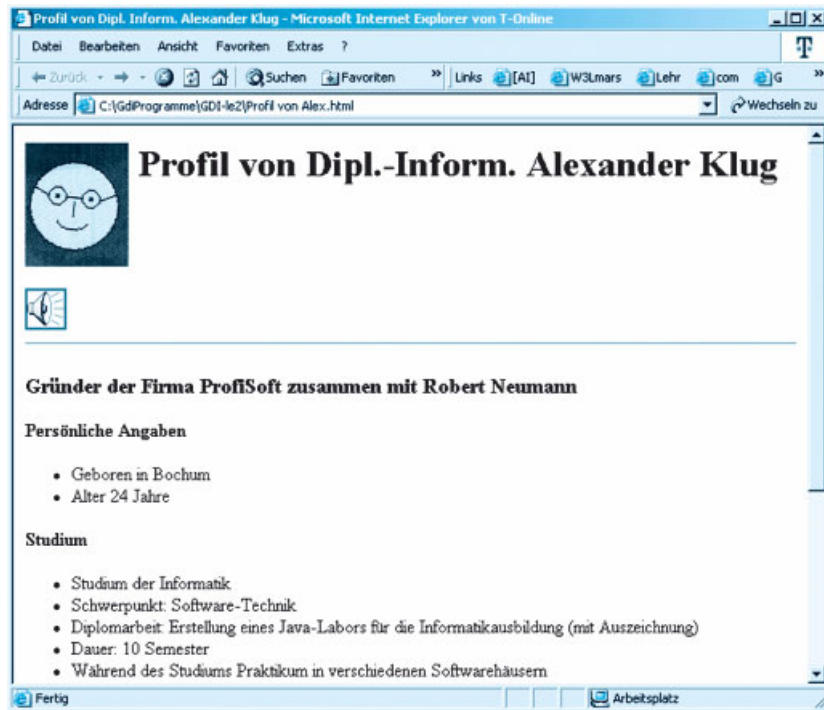
Die Firmengründer Alex und Rob erkennen schnell, dass sie ihren Bekanntheitsgrad erhöhen müssen, um genügend Aufträge zu erhalten. Als innovative Firma entscheiden sie sich dazu, auf die Dienstleistungen von ProfiSoft im Internet aufmerksam zu machen.



Abb. 1.8-1 zeigt den Teil der Website, der das Qualifikationsprofil von Alex darstellt.

## LE 2 1.8 Fallstudie: Die Firma ProfiSoft im Internet/Glossar

Abb. 1.8-1:  
Web-Präsentation  
von Alex



**ASCII** (*American Standard Code of Information Interchange*) genormter 7-Bit-Zeichensatz (128 Positionen) zur Darstellung von Ziffern, Buchstaben, Sonderzeichen und Steuerzeichen (siehe auch →Latin und →Unicode).

**Browser** →Web-Browser.

**Chats** erlauben →Internet-Benutzern interaktive Gespräche in Echtzeit.

**DNS** (*domain name system*) →*domain*-Name.

**domain-Name** Hierarchisch gegliederte, funktionsbezogene Computernamen im →Internet. Ein *Nameserver* setzt *domain*-Namen in →IP-Adressen um. Typischer Aufbau: computer.bereich.institution.land.

**E-Mail** Asynchrone Übertragung elektronischer Post, d.h. von Briefen und Nachrichten, zwischen vernetzten Computersystemen. Im →Internet wird dazu das SMTP-Übertragungsprotokoll verwendet. Der Empfänger benötigt dazu eine E-Mail-Adresse, die folgendermaßen aufgebaut ist: empfänger@smtp. server.

**Extranet** Erweiterung eines →Intranets hin zu anderen Unternehmen, Händlern, Lieferanten, die einen beschränkten Zugriff auf die Informationen des Intranets erhalten.

**ftp** (*file transfer protocol*) erlaubt im →Internet die Übertragung kompletter Dateien und Softwarepakete (ist das Pendant zur Paketpost) mit »Selbstabholung«.

**gateway** Computersystem, das Computernetze mit unterschiedlichen Übertragungsprotokollen miteinander verbindet und eine Protokollumsetzung vornimmt.

**gif** im →Web häufig benutztes Grafikformat mit maximal 256 Farben (→*jpeg*).

**Homepage** → Startseite. Oft auch als Bezeichnung für eine kleinere – meist private →Website benutzt.

**HTML** (*hypertext markup language*) Dokumentenauszeichnungssprache, die es mithilfe von Formatbefehlen (*tags*) erlaubt, inhaltliche Kategorien von Web-Dokumenten, z.B. Überschriften,

Absätze, sowie →Hyperlinks zu kennzeichnen. So ausgezeichnete Dokumente werden von →Web-Browsern interpretiert und dargestellt. Dateiendung .html bzw. .htm. →XHTML

**Hyperlinks** Verweise auf andere Dokumente; in →Web-Browsern meist farblich oder unterstrichen hervorgehoben; ein Mausklick auf einen Hyperlink bewirkt, dass zu dem Dokument, auf das verwiesen wird, verzweigt wird.

**Hypertext** Text, der Sprungmarken bzw. Verweise (→Hyperlinks) auf andere Texte enthält.

**Internet** Weltweites, dezentralisiertes, allgemein zugängliches Computernetz, in dem eine Vielzahl von Diensten angeboten und genutzt werden. Als Übertragungsprotokoll wird →TCP/IP verwendet.

**Intranet** Firmeninternes, nicht öffentliches Netz, das auf der Technik des →Internet basiert, insbesondere auf →TCP/IP.

**IP-Adresse** Eindeutig zugewiesene Adresse eines Computersystems im →Internet; besteht aus vier Bytes, durch Punkte getrennt, z.B. 134.147.80.1.

**jpeg** im →Web weit verbreitetes Bildformat, das im Gegensatz zu →gif beliebig viele Farben darstellen kann.

**Latin** genormter 8-Bit-Zeichensatz (256 Positionen), der den →ASCII-Code um 128 Positionen erweitert. Latin-1 deckt westeuropäische Sprachen ab (→Unicode).

**Leitseite** →Startseite.

**Link** →Hyperlink.

**newsgroups** »schwarze Bretter« für asynchronen Nachrichtenaustausch im →Internet.

**png** im →Web benutztes Bildformat, das sich auch für Animationen eignet (→gif, →jpeg).

**plug-ins** Software, die →Web-Browser um Zusatzfunktionen erweitern. Oft von Drittanbietern erstellt.

**Provider** stellen Zugänge zum →Internet zur Verfügung.

**search engine** →Suchmaschine.

**Startseite** (Homepage) Einstiegsseite einer Website mit →Hyperlinks zu anderen Dokumenten, wird vom →URL adressiert.

**Suchmaschine** Leistungsfähiges Computersystem, das automatisch nach bestimmten Kriterien Informationen im Internet sucht, analysiert, indiziert und z.B. in Form von Katalogen speichert und diese für Suchanfragen auswertet.

**TCP/IP** (*transmission control protocol/internet protocol*) Verfahren, wie zwischen Computersystemen im →Internet Daten übertragen werden. Anders als bei einem Telefongespräch wird keine feste Verbindung hergestellt. Statt dessen werden die Daten in Pakete von bis zu 1.500 Zeichen zerlegt, nummeriert, mit Absender- und Empfängeradresse versehen und einzeln verschickt. An den Knotenpunkten des Netzes lesen Wegplanungscomputer (*router*) die Adressen und leiten das Paket in Richtung Empfänger weiter. Das Zielcomputersystem setzt die Pakete entsprechend ihrer Nummerierung wieder zusammen.

**Telnet** erlaubt den direkten Zugriff auf andere Computersysteme im →Internet, die sich damit aus der Ferne bedienen lassen.

**Unicode** (UCS) genormter 16-Bit-Zeichensatz (65.469 Positionen), der die Schriftzeichen aller Verkehrssprachen der Welt aufnehmen soll (→ASCII und →Latin).

**URL** (*uniform resource locator*) im →Web verwendete standardisierte Darstellung von Internetadressen; Aufbau: protokoll://domain-Name/Dokument-pfad.

**W3** Kurzform für →WWW.

**Web** Kurzform für →World Wide Web.

**Web-Browser** Software, über die Benutzer die Dienstleistungen des Internet, insbesondere des →WWW, in Anspruch nehmen können. Durch Angabe der →URL wird das Computersystem, das die jeweilige Dienstleistung anbietet, eindeutig adressiert.

**Website** Das komplette Online-Angebot eines Anbieters (Privatperson, Organisation, Unternehmen) unter einer →URL, das üblicherweise mit der →Startseite beginnt. Eine Website besteht aus mindestens einer – in der Regel aber mehreren →Webseiten, die über →Links untereinander verknüpft sind. Die einzelnen Webseiten müssen sich dabei

## LE 2 Glossar/Zusammenhänge

nicht zwangsläufig auf nur einem Server befinden.

**Webseite** Eine Seite, die in einem →Web-Browser angezeigt wird. In der Regel handelt es sich um ein HTML-Dokument, das Texte, Medien (Bilder, Video, Audio), →Links und Programme enthalten kann. Bestandteil einer →Website.

**World Wide Web** →Web.

**WWW** (*World Wide Web*) Informationssystem im →Internet, das auf der →Hypertext-Technik basiert; ermöglicht außerdem den Zugriff auf die anderen Internetdienste.

**XHTML** (*eXtensible HyperText Markup Language*) Dokumentenbeschreibungssprache, die die Befehle von →HTML 4 enthält und der Syntax von XML (*eXtensible Markup Language*) folgt. Strenger als HTML. Ein XHTML-Dokument kann daraufhin geprüft werden, ob es wohlgeformt (*well-formed*) und gültig (*valid*) ist. Wohlgeformt bedeutet, dass sich ein Dokument an die syntaktischen Regeln von XHTML hält. Ein Dokument ist gültig, wenn seine Struktur den Vorgaben eines Dokumententyps entspricht. Im Jahr 2000 als W3C-Standard (*World Wide Web Consortium*) verabschiedet. Nachfolger von HTML.

Lokale Computernetze werden durch das **Internet** weltweit verknüpft (Netz der Netze). Im Internet werden verschiedene Dienste angeboten, unter anderem:



Internetdienste

- **Web** (*World Wide Web*), auch **W3** oder **WWW** genannt: ein hypertext-basiertes Informationssystem,
- **E-Mail**: asynchrone, elektronische Post für Briefe und Nachrichten,
- **ftp**: Transportdienst für Dateien (elektronische »Paketpost«),
- **Newsgroup**: asynchrone Computer-Konferenzen und »schwarze Bretter«,
- **Chats**: interaktives »Plaudern« in Echtzeit,
- **Telnet**: Fernbedienen eines entfernten Computersystems.

Um die Web-Dienstleistungen in Anspruch nehmen zu können, wird ein **Web-Browser**, kurz **Browser** genannt, benötigt. Er ist in der Lage, **HTML**-Dokumente und **XHTML**-Dokumente zu interpretieren und anzuzeigen. Ein XHTML-Dokument verwendet die **Hypertext**-Technik, um mithilfe von **Hyperlinks** auf andere Dokumente zu verweisen. Die Einstiegsseite einer **Website** heißt **Startseite**, **Leitseite** oder **Homepage**. Von dort wird auf andere **Webseiten** verzweigt. Um die Übertragungszeiten im Netz zu reduzieren, werden in Web-Dokumenten besondere Grafik- und Bildformate verwendet. Verwendet werden heute **gif**, **jpeg** und **png**. Damit Browser zusätzliche Formate anzeigen können, werden Browser über **plug-ins** erweitert.

Web

Über Browser können nicht nur die Web-Dienste im Internet, sondern auch alle anderen in Anspruch genommen werden, d.h., Browser stellen eine einheitliche Benutzungsoberfläche für alle Dienste zur Verfügung. Damit ein bestimmter Dienst auf dem jeweils gewünschten Computersystem angesprochen werden kann, wird das **URL**-Adressschema verwendet:

Protokoll des gewünschten Dienstes://domain-Name/Dokumentpfad

Der **domain-Name** ist nach dem *domain name system (DNS)* strukturiert: computer.bereich.institution.land

Über einen *Nameserver* wird jedem *domain*-Namen eine **IP-Adresse** zugeordnet, die aus vier Bytes besteht. Für den E-Mail-Dienst wird zusätzlich der Empfängername benötigt. Der Adressaufbau sieht daher folgendermaßen aus: empfänger@smtp.server.

Viele Dienste im Internet verwenden zur Informationsdarstellung noch den **ASCII-Code**, im Web kann bereits der **Latin-Code** benutzt werden. Der **Unicode** wird bisher nur im Betriebssystem Windows und in der Programmiersprache Java benutzt.

Um Informationen im Internet zu finden, können verschiedene Suchstrategien angewandt werden. **Suchmaschinen (search engines)** können bei der Suche helfen.

Die Technik des Internet beruht auf dem **TCP/IP**-Übertragungsverfahren, bei dem Daten, in Pakete aufgeteilt, asynchron über das Netz – unter Umständen auf verschiedenen Wegen – transportiert werden. Netze, die ein anderes Übertragungsprotokoll verwenden, werden über **gateways** an das Internet angeschlossen. **Intranets** und **Extranets** benutzen dieselbe Technik wie das Internet, sind aber *nicht* öffentlich zugänglich.

Der Anschluss an das Internet erfolgt über einen **Provider**.

Die rasante Entwicklung des Internet in den letzten Jahren wird es in Zukunft zu einem Massenmedium wie heute Fernsehen und Telefon werden lassen.



/Balzert, Balzert, Krengel 04/

Balzert, Helmut; Balzert, Heide; Krengel, A., *Das Internet – Beruflich und privat effizient und sicher nutzen*, Herdecke-Dortmund: W3L-Verlag 2004  
Umfassende Einführung in die Internet-Welt.

Literatur

/Balzert 03a/

Balzert, Helmut, *HTML, XHTML & CSS für Einsteiger – Statische Websites systematisch erstellen*, Herdecke-Dortmund: W3L-Verlag 2003

Zitierte Literatur

/Balzert 03b/

Balzert Helmut, *HTML & XHTML (Quick Reference Map)*, Herdecke-Dortmund: W3L-Verlag 2003

/Balzert 04/

Balzert Heide, *Webdesign & Web-Ergonomie – Websites professionell gestalten*, Herdecke-Dortmund: W3L-Verlag 2004

/IEEE Spectrum 97/

*Communications*, in: IEEE Spektrum, Jan. 1997, p. 27

/Kühnert 97/

Kühnert H., *Die durchgedrehte Ligatur*, in: Die ZEIT, 7.3.1997

/Meissner 96/

Meissner R., *Mischobst – Grafikformate für Web-Dokumente*, in: c't, Heft 7, 1996, S. 206–210

/Schneider 95/

Schneider G., *Eine Einführung in das Internet*, in: Informatik-Spektrum, 18, 1995, S. 263–271

/Zimmer 98/

Zimmer D. E., *Deutsch und anders*, Hamburg: Rowohlt-Verlag 1998

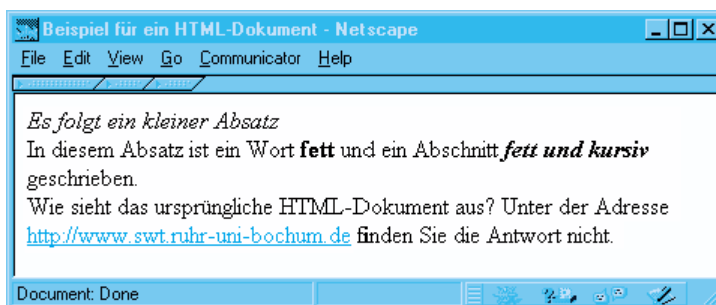
## LE 2 Aufgaben

- Analytische Aufgaben  
Klausur-Aufgabe  
5 Minuten
- 1 *Lernziel: Multimediale XHTML-Dokumente bezogen auf die Syntax analysieren können.*  
Ergänzen bzw. korrigieren Sie den nachfolgenden XHTML-Code, sodass er syntaktisch korrekt ist.

```
<html>
<head>
<title>
<a1>Top-Seite der Ruhr-Universität Bochum<H1>
</endtitle>
<body>
<bold>Diese Zeile soll fett sein</bold>
Diese Zeile ist nicht mehr fett
Diese Zeile soll kursiv sein !
</endbody>
</end>
```

- Konstruktive Aufgaben  
Muss-Aufgabe  
20 Minuten
- 2 *Lernziel: Mithilfe der angegebenen Befehle multimediale XHTML-Dokumente erstellen und auf Web-Servern bereitstellen können.*  
Betrachten Sie die in Abb. 1.8-2 wiedergegebene XHTML-Seite. Versuchen Sie, ein XHTML-Dokument zu erstellen, das genau diese Darstellung erzeugt. Überprüfen Sie Ihr Ergebnis, indem Sie Ihr Dokument in einen Web-Browser laden.

Abb. 1.8-2:  
Zu erstellende  
XHTML-Seite



- Muss-Aufgabe  
45 Minuten
- 3 *Lernziel: Mithilfe der angegebenen Befehle multimediale XHTML-Dokumente erstellen und auf Web-Servern bereitstellen können.*  
Erstellen Sie eine Startseite mit Angaben zu Ihrer Person, einem Bild und einer gesprochenen Information. Machen Sie anschließend Ihre Startseite im Web bekannt. Orientieren Sie sich an der Fallstudie von Kapitel 1.8. Verwenden Sie die Informationen aus Aufgabe 2 der Lehrinheit 1.
- 4 *Lernziel: Anhand der beschriebenen Suchstrategien systematisch im Internet unter Einsatz von Suchmaschinen Informationen recherchieren können.*
- a Versuchen Sie mithilfe eines Web-Suchdienstes und eines Personen-Suchdienstes herauszufinden, welche Informationen über Personen mit Ihrem Namen im Internet verfügbar sind.
  - b Recherchieren Sie nach deutsch- und englischsprachigen Informationen zur Internet-Netiquette und fügen Sie die gefundenen Verweise in Ihre Homepage ein.
  - c Recherchieren Sie nach Informationen über die Gestaltung und Erstellung von XHTML-Seiten und fügen Sie die gefundenen Verweise in Ihre Homepage ein.
  - d Finden Sie alle Webseiten heraus, von denen auf Ihre Homepage verwiesen wird.
- Hinweis Weitere Aufgaben befinden sich auf der CD-ROM.

