Hypertexte im Internet –

Die Hypertext Markup Language (HTML) und das World Wide Web (WWW)

Vortrag im Informatik-Seminar „Hypertexte“ im Sommersemester 1994

Kai Garlipp, 9. Mai 1994
Email: garlipp@informatik.uni-rostock.de
WWW: http://www-dbis.informatik.uni-rostock.de/Kai/

Inhalt

1 Einführung ____ *3*
1.1 NetNews ____ **3**
1.2 Gopher ____ **3**
1.3 Wide Area Information Server (WAIS) ____ **5**
1.4 World Wide Web (WWW) ____ **5**
2 Die Hypertext Markup Language (HTML) ____ *6*
2.1 Die Struktur der HTML ____ **7**
2.2 HTML Sprachelemente ____ **8**
2.2.1 HEAD und BODY ____ 9
2.2.2 TITLE ____ 9
2.2.3 Überschriften, H1 – H6 ____ 9
2.2.4 Zeichenhervorhebung ____ 10
2.2.5 Absätze ____ 11
2.2.6 Darstellung von Listen ____ 12
2.2.7 Anker ____ 13
2.2.8 Einbinden von Grafiken und Bildern (IMG) ____ 14
2.2.9 Erstellen von Tabellen (Table) ____ 15
2.3 Universal Resource Locator (URL) ____ **17**
2.4 Style Guide ____ **18**
2.5 Erweiterungen von HTML zu HTML+ ____ **19**
3 HTML-Beispiele ____ *19*
4 Literatur ____ *20*
4.1 NetNews ____ **20**
4.2 Gopher ____ **20**
4.3 WAIS ____ **20**
4.4 WWW und HTML ____ **20**

1 Einführung

Das Internet ist das älteste und zur Zeit weltweit größte Computernetzwerk mit ca. zwei Millionen angeschlossenen Rechnern und mehr als zehn Millionen Benutzern (Mitte 1993). Von praktischem Nutzen sind vor allem die vielfältigen Netzwerkdienste, die eine Vielzahl von allgemeinzugänglichen Kommunikationsmitteln ermöglichen. So existieren neben den Basisdiensten Telnet, FTP und Mail noch etwa 100 weitere verschiedene Protokolle der TCP/IP-Familie, von denen einige, die für den Umgang im World Wide Web (WWW) wichtig sind, hier beschrieben werden.

1.1 NetNews

Die Grundlagen für die NetNews wurden seit 1979 von Steve Bellovin, Jim Ellis, Tom Truscott und Steve Daniel an der Duke University in den USA geschaffen. Ihr Ziel war der Informationsaustausch innerhalb der Unix-Gemeinde via Computer. Um diese Nachrichtenschwemme zu koordinieren, finden die Diskussionen in thematisch gegliederten News-Gruppen statt.

Die Namensgebung der News-Gruppen ist hierarchisch organisiert: jeder Name setzt sich aus mehreren durch Punkte getrennten Teilnamenskürzeln zusammen, wobei der linke Teilname stets die übergeordnete Kategorie beschreibt. Beispielsweise liest sich die News-Gruppe ‘comp.infosystems.www’ von links nach rechts wobei der linke Teil (‘comp’) die höchste Kategorie bzw. Hierarchiestufe beschreibt. ‘comp’ steht für Computer, ‘infosystems’ bezeichnet, daß das Thema dieser Gruppe Informationssysteme sind. ‘www’ schließlich sagt aus, daß Fragen zum World Wide Web diskutiert werden[1].

Den News-Gruppen werden die einzelnen Artikel zugeordnet. Jeder Artikel erhält von dem Server, auf dem er gespeichert ist, eine eindeutige Nummer, z.B. 9311120229.AA24657@expo.x.org. Die Numerierung ist fortlaufend und separat für jede Gruppe.

Um an den NetNews teilzunehmen, wird ein Zugang via Network News Transport Protokoll (NNTP) zu einem News-Server benötigt. An UNIX-Systemen wird der News-Server mittels der Umgebungsvariable `NNTPSERVER` eingestellt, z.B. um den News-Server der Freien Universität Berlin zu benutzen, ist `setenv NNTPSERVER news.fu-berlin.de` anzugeben.

1.2 Gopher

Eigentlich ist Gopher eine Taschenratte. Darüber hinaus verbirgt sich dahinter das Maskottchen des Bundesstaates Minnesota, der gelegentlich auch mit ‘The Gopher State’ bezeichnet wird. Die Ratte ist auch das Maskottchen der Universität von Minnesota.

---

[1] Die Newsgruppe *comp.infosystems.www* wurde Anfang September 1994 ersetzt durch die neuen Newsgruppen *comp.infosystems.www.users*, *comp.infosystems.www.providers* und *comp.infosystems.www.misc*. (24. Oktober 1994)

Die erste Gopher Software und das dazugehörige Gopher-Protokoll wird als verteiltes Informationssystem (Campus-Wide Information System CWIS) seit 1991 an der University of Minnesota unter der Leitung von Dave Johnson, Mark McCahill, Paul Lindner, Bob Alberti, Daniel Torrey, Helge Tonsky und Farhad Anklesia entwickelt. Es sollten hauptsächlich Informationen, Neuigkeiten, Ankündigungen und sonstige Informationen in einfacher Textform elektronisch für Studierende und Universitätspersonal zur Verfügung gestellt werden. Dabei sollen Daten der einzelnen Departements dezentral zu verwalten, aber zusammen mit den Informationen der anderen Abteilungen zentral zugänglich sein.

Die Fähigkeiten des von Gopher sind inzwischen erheblich erweitert worden (z.B. die Erweiterung des Gopherprotokolls zu Gopher+), der Grundgedanke ist aber erhalten geblieben: Gopher stellt Internet-Dienste für den Benutzer zentral zur Verfügung, wobei die geographische Verteilung der Informationen und die unterliegenden Abläufe des Gopherprotokolls verdeckt bleiben.

Um ein Gopherdokument von einem Server abrufen zu können, benötigt der Gopher-Client den Hostnamen des Servers, die Angabe der Nummer des Kommunikationsports (meistens Port 70), eine Angabe, wo sich das gewünschte Dokument auf diesem Server befindet (Selektor, vergleichbar mit einem Pfadnamen) sowie den Typ des anzuzeigenden Objektes (z.B. Textdatei, Unterverzeichnis, Bilddatei, Telnet-Sitzung). Ein Beispiel dafür sei der Eintrag für den Ordner ‘Fachbereich Informatik’ vom Gopher-Server des Fachbereichs:

Name=Fachbereich Informatik
Host=ceylon.informatik.uni-rostock.de
Port= 70
Type=1 (Unterverzeichnis)
Path=1/Fachbereich Informatik

Der normale Gopher-Benutzer braucht diese internen Angaben über den Hostnamen und den Pfad aber nicht zu wissen, in seinem Gophermenü wird nur der Name des Antrages angezeigt.

Die folgende (lustige) Definition des Gophers kann auf dem Gopher-Server der University of Minnesota gefunden werden:

Gopher n. **1.** Any of various short tailed, burrowing mammals of the family Geomyidae, of North America. **2.** (Amer. colloq.) Native or inhabitant of Minnesota: the Gopher State. **3.** (Amer. colloq.) One who runs errands, does odd-jobs, fetches or delivers documents for office staff. **4.** (Computer tech.) software following a simple protocol for burrowing through a TCP/IP internet.

Gopher+ n. 1. Hardier strains of mammals of the family Geomyidae. **2.** (Amer. colloq.) Native or inhabitant of Minnesota, the Gopher state, in full winter regalia (see PARKA). **3.** (Amer. colloq.) Executive secretary. **4.** (computer tech.) Software following a simple protocol for burrowing through a TCP/IP internet, made more powerful by simple enhancements (see CREEPING FEATURISM).

TurboGopher n. **1.** A small rodent with a turbocharger strapped on its back to increase its speed and ferocity. **2.** (Amer. colloq.) Native or inhabitant of Minnesota after consuming three double espressos. **3.** (Amer. colloq.) An Olympic sprinter who runs errands, does odd-jobs, fetches or delivers documents for office staff. **4.** (computer tech.) Speed-optimized Macintosh software following a simple protocol for burrowing through a TCP/IP internet; network speed is achieved by using turbocharged software; incoming bits spin the turbine that pumps out the outgoing bits.

Um ein Gopher-Dokument, von dem der Menüeintrag (zumindest teilweise) bekannt ist, zu finden, bedient man sich eines Veronica-Servers. Nach Eingabe eines Suchbegriffs erhält der Benutzer die dem Server bekannten Dokumente in Form eines Gophermenüs zurück und er kann die einzelnen Dokumente anschließend auswählen und betrachten.

1.3 Wide Area Information Server (WAIS)

Schon seit langem stehen auf vielen Rechnern im Internet Informationsdienste zur Verfügung, z.B. Online-Bibliothekskataloge, die über eine Telnet-Sitzung erreicht werden können. Hinter solchen öffentlichen Logins auf anderen Rechners verbergen sich meist selbstgeschriebene Menüsysteme, die nicht standardisiert sind.

Wide Area Information Server (WAIS), entwickelt von den Thinking Machines Corporation, erleichtern die Abfrage an solche Datenbanken, indem sie deren unterschiedliche Retrieval-Sprachen umgehen und dem Benutzer ein einfaches und einheitliches Interface präsentieren. Nach der Eingabe eines Suchbegriffes (z.B. ein (Zeitschriften)artikel, ein Buch, eine abstrakte oder spezielle Information) zeigt WAIS ein Menü der ihm bekannten Dokumente im Netz an, die die gesuchte Information aufweisen. Durch Anwahl eines Menüpunktes erfolgt die Auswahl eines Dokuments, das der Benutzer anschließend online einsehen kann. Informationen werden also nach inhaltlichen Kriterien gesucht und nicht auf Grund ihrer Dateinamen (wie bei den Archie-Servern zur Bestimmung eines Dokuments auf FTP-Servern) oder ihrer Menübezeichnung (bei Veronica-Servern im Gopherspace) gefunden.

Anfragen an eine WAIS Datenbank können nicht nur von speziellen WAIS-Clients aus getätigt werden, sondern auch aus dem Gopher und WWW heraus erfolgen.

1.4 World Wide Web (WWW)

Mittlerweile ist die Menge der im Internet angebotenen Informationen so groß, daß eine Aufzählung nur unvollständig seien kann. Beispiele sind Auskunftssysteme von Universitäten, Bibliotheken, Wetterdiensten, Presseagenturen, Regierungen und deren Behörden, Recherchemöglichkeit in Online-Zeitschriften und Enzyklopädien, Interessierte können auch historischen Reden oder Volksmusik lauschen, sich Bilder einer Kunstgalerie oder auch Filme ansehen. Auch immer mehr (nicht Computer-) Unternehmen entscheiden sich zu einem Internetanschluß, um ihre Informationen, z.B. über Ausschreibungen und aktuelle Projekte, einem breiten Publikum präsentieren zu können.

Diese Beispiele machen deutlich, wie komplex das Internet ist. 1989 leitete deshalb eine Gruppe von Physikern am Kernforschungszentrum CERN (Centre Europeén de Recherches Nucléairs) in Genf ein Softwareprojekt in die Wege, das als globales, interaktives Informationssystem für die Hochenergiephysik dienen sollte. Doch schon recht bald war klar, daß ein solches Informationssystem nicht nur für einen begrenzten Ausschnitt, sondern auch für das gesamte Internet von großem Nutzen wäre. Denn mit seiner Hilfe ließen sich im Netz die Informationen unter einer einheitlichen Schnittstelle integrieren. Das Ergebnis dieser Entwicklungsarbeiten ist heute unter den Namen World Wide Web (WWW oder kurz W3) bekannt. Es ist aber mehr eine Initiative als ein Produkt. Ziel ist es, mit Hilfe von Hypermedia-Informationssystemen universellen Zugang zu global verteilten Dokumenten zu gewähren. Hypermedia bezeichnet in Abgrenzung von Hypertext nicht nur schriftliches Material, sondern auch andere Medien wie Sprache, Musik, Grafiken, Bilder und Filme.

Um ein Dokument von einem Hypertextserver anzufordern und dieses Dokument, wenn es dort vorhanden ist, anschließend zum WWW-Client zu transferieren, wird das Hypertext Transfer Protokoll (HTTP) benutzt, das in seinen Grundzügen wie folgt funktioniert:

1. Der Client baut eine Verbindung zum Server auf, der standardmäßige Kommunikationsport ist dabei der Port 80.
2. Mit Kommando `GET <Lokator>` wird dann das mit `<Lokator>` bezeichnete Dokument vom Server angefordert.
3. Der Server sendet des angeforderte Dokument oder eine Fehlermeldung an den Client und beendet die Kommunikation.

Einige weitere Kommandos dienen außerdem dem Senden von Dokumenten an einen Server oder aber der Suche von Strings in den Textdokumenten. Der Benutzer von WWW-Clients kommt aber mit diesem Protokoll und seinen Befehlen nicht in Kontakt.

WWW-Clients sind heute für die unterschiedlichsten Plattformen und Systeme im Internet frei erhältlich. Am verbreitetsten ist dabei der Mosaic-Browser vom NCSA (National Center for Supercomputing Applications), der sowohl für X11, MS-Windows und den Apple Macintosh verfügbar ist, für X11 existieren weiterhin Viola und tkWWW und für MS-Windows Cello[2]. Benutzer von ASCII-Terminals können Hypertexte mit dem WWW-Client Lynx, der seit kurzem auch als DOS-Version erhältlich ist, betrachten.

Zwischen mehreren frei verfügbaren WWW-Servern für UNIX-Systeme kann ein Informationsanbieter wählen[3]. Die beiden wichtigsten sind die Server des CERN und des NCSA, daneben gibt es den Perl geschrieben Plexus und den GN-Server, der gleich zwei Protokolle versteht und sowohl Gopher als auch WWW unterstützt.

2 Die Hypertext Markup Language (HTML)

WWW-Autoren erstellen ihre Hypertextdokumente in HTML. Die Sprache HTML ist gegenwärtig in einem Internet-Draft definiert, also einem Internet-Standard in Vorbereitung. HTML ist eine Standard Generalized Markup Language (SGML), d.h. es wird nur die logische Struktur eines Dokuments wie Absätze, Überschriften, Aufzählungslisten, Texthervorhebungen beschrieben. Layout-Informationen dagegen entfallen in HTML-Dokumenten. Der Autor legt lediglich den Text fest, z.B. eine Überschrift, hat aber keinen Einfluß darauf, in welcher Schriftart und Größe sie beim Benutzer erscheint. Die Darstellung hängt vielmehr von den technischen Möglichkeiten der Clients ab. Schreibt ein HTML-Autor in seinem Text etwa ‘<STRONG> halbfett </STRONG>‘, so mag das auf seinem Bildschirm durchaus in halbfetter Schrift erscheinen, auf andern Clients aber auch in Kapitälchen oder wie an ASCII-Terminals gar nicht hervorgehoben. Bei der Beschreibung der einzelnen Sprachelemente wird daher in den folgenden Kapiteln von der standardmäßigen Darstellung laut Internet Draft ausgegangen. Bei den HTML-Sprachelementen wird desweiteren zwischen drei verschiedenen Stadien der Implementierung in den HTML-Parsern unterschieden:

---

[2] Starke Verbreitung fand in letzter Zeit auch der WWW-Browser Netscape, der für MS-Windows, die verbreitetsten UNIX-Systeme und Macintosh verfügbar ist. (25. Mai 1995)

[3] Eine Liste aller verfügbarer Server und Clients wird auf dem WWW-Server der WWW-Organisation (*http:/www.w3.org /*) geführt. (24. Juni 1994/25. Mai 1995)

Mainstream: Alle Parser müssen diese Sprachelemente erkennen.

Extra: Diese HTML-Standardsprachelemente dürfen durch die Parser ignoriert werden. Es ist auch erlaubt, diese Elemente so zu ignorieren, als wären sie nicht vorhanden (z.B. bei EM).

Obsolete (veraltet): Diese Sprachelemente gehören nicht mehr zur Standard-HTML. Die Parser sollten diese Sprachelemente implementieren um eine Abwärtskompatibilität zu vorherigen Versionen des HTML-Standards zu gewährleisten.

Dieser Status wird bei der Beschreibung der einzelnen Sprachelemente in den folgenden Kapiteln angegeben, erfolgt kein Hinweis, ist der Status des Elementes Mainstream.

Die Erstellung von HTML-Dokumenten gestaltet sich leider noch ziemlich mühsam. Der Autor muß die Dokumente mit Hilfe eines normalen Texteditors direkt in HTML erstellen[4]. Er muß deshalb alle relevanten Tags bzw. Kommandos kennen, WYSIWYG ist (noch) nicht möglich. Es gibt aber auch Werkzeuge, die den Autor beim Setzen der HTML Tags ein wenig unterstützen, z.B. HTML-Write für MS-Windows. Ein anderer Ansatz zur Erstellung von HTML-Dokumenten beruht auf der automatischen Konvertierung aus anderen Dokumentformaten nach HTML. Es existieren Tools, die beispielsweise LaTeX-Dokumente (*latex2html*) und MS Word Rich Text Format Dokumente (RTF, *rtftohtml*) nach HTML konvertieren.

2.1 Die Struktur der HTML

Eine HTML-Instanz ist wie gewöhnlicher Text, mit der Ausnahme, daß einige Zeichen als Markierung interpretiert werden. Diese Markierungen sind die Grundstruktur des Dokuments.

Die Instanz repräsentiert eine Hierarchie von Sprachelementen. Jedes Sprachelement hat einen Namen, einige haben Attribute und einige einen Inhalt. Die meisten Elemente werden im Dokument repräsentiert als Startzeichen (start tag), die den Namen und die Attribute enthalten, gefolgt durch den Inhalt und einem abschließenden das Ende-Zeichen (end tag). Startzeichen werden begrenzt durch < und >, die Endezeichen durch </ und >. Bei den Namen wird nicht zwischen Groß- und Kleinschreibung unterschieden. Ein Beispiel dafür ist:

[4]Es sind auch WYSIWYG-Editoren für die Erstellung von HTML-Texten verfügbar. SoftQuard HoTMetal ist für SPARCstations und für MS-Windows frei erhältlich, z.B. auf *ftp://ftp.germany.eu.net/pub/infosystems/www/ncsa/Mosaic/contrib/SoftQuard* (24. Oktober 1994) Außerdem sind Formatvorlagen für Textverarbeitungssysteme, z.B. für Word for Windows 6.0, die das Erstellen von HTML-Dokumenten ermöglichen, verfügbar. Ein Beispiel ist der „Internet Assistant for Word 6.0“ von Microsoft (http://www.microsoft.com/) der auch das Web-Browsen direkt aus Word heraus ermöglicht.

```
<HTML>
      <TITLE>
            Ein einfaches HTML Beispiel
      </TITLE>
      <H1>
            Ein Beispiel
      </H1>
      Hier folgt ein Absatz.
      <P>
      <UL>
      <LI>Element eins hat einen
            <A NAME="anchor">
                  Anchor.
            </A>
      <LI>Hier ist Element zwei.
      </UL>
</HTML>
```

Einige Elemente sind leer, z.B. P und LI; sie haben keinen Inhalt und werden nur als Startzeichen repräsentiert.

Das HTML-Dokument besteht also aus Fließtext mit eingestreuten Strukturelementen. Keine Rolle spielen die im HTML-Quelltext vorhandenen Zeilenumbrüche und vorhandene Leerzeilen werden ignoriert. Der Zeichensatz ist ISO 8859-1, das heißt, Umlaute und andere Sonderzeichen bereiten den HTML-Autor keinerlei Probleme, solange man mit den 256 Zeichen des ISO-Standards auskommt.

Einige Zeichen verlangen aber dennoch eine Sonderbehandlung, z.B. wenn in einem Dokument die spitzen Klammern der HTML-Tags erscheinen sollen. Für diese Zeichen existiert eine Ersatzdarstellung. Das ‘<‘-Zeichen wird im HTML-Quelltext als ‘<’ und das ‘>‘-Zeichen als ‘>’ eingegeben. Das bedingt gleichzeitig, daß das ‘&’-Zeichen als ‘&’ darzustellen ist. Auch andere in der ISO-Norm definierte Sonderzeichen und typographische Symbole lassen sich auf diese Art darstellen. So kann die Darstellung des ‘ä’ als ‘ä’ erfolgen, was beispielsweise wichtig ist, wenn dem HTML-Autor keine deutsche Tastatur zur Verfügung steht. Eine Aufzählung aller auf diese Art darstellbarer Zeichen befindet sich z.B. im HTML-Internet-Draft auf Seite 23 ff.

Ein Leerzeichen muß vom WWW-Client als leerer Bereich, mehrfache Leerzeichen als größerer leerer Bereich, angezeigt werden. Die Anzeige des TAB-Zeichens ist nicht definiert, es sollte daher auch nicht verwendet werden.

In einem HTML-Text können auch Kommentare enthalten sein, die durch den Parser ignoriert werden. Kommentare werden in <!-- und --> eingeschlossen.

2.2 HTML Sprachelemente

In diesem Kapitel werden einige ausgewählte und für die Gestaltung einfacher Dokumente wichtige HTML-Sprachelemente beschrieben. Eine ausführliche Beschreibung aller Sprachelemente befindet sich im bereits erwähnten Internet-Draft.

2.2.1 HEAD und BODY

HTML-Dokumente sollten (es ist aber nicht unbedingt nötig) ein einleitendes HEAD Element besitzen, das von ein BODY Element gefolgt wird.

Das HEAD Element enthält alle allgemeinen Informationen über das Dokument. Es enthält aber keinen Text, der zum eigentlichen Dokument gehört, dieser befindet sich im BODY Element. Im HEAD Element sind nur einige weitere Elemente erlaubt, z.B. TITLE.

Das BODY Element enthält alle Informationen, die Teil des eigentlichen Dokuments sind, im Gegensatz zum den Informationen über das Dokument im HEAD.

2.2.2 TITLE

Der Titel eines Dokuments wird im TITLE Element spezifiziert. Das TITLE Element sollte im HEAD Element eines Dokuments erscheinen. Es darf in einem Dokument nur ein Titel enthalten sein. Der Titel sollte den Inhalt des Dokuments in einem ziemlich allgemeinen Zusammenhang identifizieren. Der Titel ist kein Bestandteil des Textes eines Dokuments, aber er ist ein Eigentum des gesamten Dokuments. Er darf keine Anchors, Absatzzeichen oder Texthervorhebungen enthalten.

Der Titel wird genutzt, um das jeweilige Dokument in der Dokumenten History-Liste identifizieren zu können oder wird in der Fensternamenszeile des WWW-Clients angezeigt. Darum sollte der Titel auch nicht länger als 64 Zeichen sein; der Titel darf zwar auch länger sein, aber einige Clients könnten den Titel auf die maximale Titelzeilenlänge von 64 Zeichen beschneiden.

2.2.3 Überschriften, H1 – H6

Die HTML unterstützt sechs verschiedene Ebenen zur Gestaltung von Überschriften. Eine Überschrift bewirkt eine Änderung des Darstellungsschrift und einen Absatz vor und nach der Überschrift. Die Angabe von weiteren Zeichenhervorhebungen und Absätzen sind zur Darstellung einer Überschrift in der HTML nicht erforderlich.

H1 ist die höchste Überschriftenebene und es wird empfohlen, sie für die Hauptüberschrift eines Dokuments zu verwenden. Die Hauptüberschrift sollte ein Dokument spezieller charakterisieren, und für Leser geeignet sein, die schon ähnliche Informationen gelesen haben; im Gegensatz zur Titelzeile, die das Dokument in einem weiteren Zusammenhang identifiziert.

Die einzelnen Überschriftenelemente sind:

```
<H1>, <H2>, <H3>, <H4>, <H5>, <H6>
```

Ein Beispiel dafür ist:

```
<H1>Dies ist eine Überschrift</H1>
Hier folgt etwas Text
```

```
<H2>Eine Zwischenüberschrift</H2>
Und noch mehr Text.
```

Der besseren Gestaltung von Texten, sollte einer Überschrift direkt keine Überschrift folgen, die zwei oder mehr Ebenen tiefer liegt, z.B. eine H1-Überschrift soll keine H3-Überschrift folgen.

Die einzelnen Überschriftenebenen werden typischerweise folgendermaßen dargestellt:

H1 Sehr große, fette Schrift, zentriert; eine oder zwei Zeilen Abstand zwischen der Überschrift und dem folgenden Text; bei Ausdrucken sollte eine neue Seite beginnen[5]

H2 Große, fette Schrift, linksbündig, nicht eingerückt; eine oder zwei Zeilen Abstand nach oben und unten.

H3 Große, kursive Schrift, etwas vom linken Rand eingerückt; eine oder zwei Zeilen Abstand nach oben und unten.

H4 Normale, fette Schrift, weiter eingerückt als H3; eine Zeile Abstand nach oben und unten

H5 Normale, kursive Schrift, so weit eingerückt wie H4; eine Zeile Abstand nach oben

H6 Fett, so weit eingerückt der normale Text, eine Zeile Abstand nach oben.

2.2.4 Zeichenhervorhebung

Der Status der Zeichenhervorhebung ist extra, d.h. die Parser können diese Sprachelemente ignorieren, der Inhalt der Elemente wird aber trotzdem dargestellt.

Die Elemente der Zeichenhervorhebung erlauben es, Teile von Texten auf verschiedenste Weise zu formatieren. Sie bewirken keinen Zeilenvorschub und können beliebig innerhalb von Textabschnitten auftreten. Zu allen Hervorhebungselementen gehört das jeweilige Endezeichen, wie in

```
Das ist <EM>hervorgehobener</EM> Text.
```

Da einige Browser diese Hervorhebungselemente ignorieren, sollte ein HTML-Autor Formulierungen wie ‘Bitte beachten sie den fettgedruckten Abschnitt’ vermeiden.

In der HTML wird zwischen physischen und logischen Stylingelementen unterschieden. Die physischen Elemente sind:

---

[5] Derzeit zentrieren die meisten Browser H1-Überschriften nicht (auch wenn es so im Internet-Draft steht). Deshalb hier an dieser Stelle der Tip für zentrierte Überschriften: <H1 ALIGN=CENTER> (04. Juni 1995)

TT	Fester Zeichenabstand, wie bei der Schreibmaschine (TT steht für Teletype)
B	Fett (bold)
I	Kursiv (italic)
U	Unterstrichen

Die logischen Elemente sind:

EM	Hervorgehoben(emphasis), typischerweise kursiv
STRONG	Stärker hervorgehoben, typischerweise fett
CODE	Ein Code-Beispiel, üblicherweise nichtproportional
SAMP	Eine Buchstabenfolge
KBD	In einer Bedienungsanleitung vom Nutzer eingegebener Text
VAR	Ein Variablenname
DFN	Ein zu definierender Begriff, typischerweise fett oder kursiv
CITE	Ein Zitat, typischerweise kursiv

Ein Beispiel:

```
Dieser Text enthält ein <em>hervorgehobenes</em> Wort.
<strong>Erwartet nicht,</strong> das hier etwas in kursiv steht.
Hier wurde das <code>EM</code> Element gebraucht.
<cite>Moby Dick</cite> ist ein Buchtitel.
```

2.2.5 Absätze

Mit dem leeren Sprachelement P wird ein Absatz gekennzeichnet. Es wird zur Trennung zweier Textteile benötigt, die sonst zusammen erscheinen würden.

Das <P> Element wird nicht benötigt, um Leerzeilen um Überschriften, Listen oder Adressen herum zu gestalten. Auch innerhalb von PRE Elementen (vorformartierter Text) sollte das <P> nicht verwendet werden. Ein Beispiel für den richtigen Gebrauch:

```
<h1>What to do</h1>
This is a one paragraph.<p>This is a second.
<P>
This is a third.
```

Ein Beispiel für den falschen Gebrauch, keines der Absatzzeichen in diesem Beispiel sollte verwendet werden:

```
<h1><P>What not to do</h1>
<p>I found that on my XYZ browser it looked prettier to
me if I put some paragraph marks
<p>
<ul><p><li>Around lists, and
<li>After headings.
</ul>
<p>
```

2.2.6 Darstellung von Listen

Die HTML definiert verschiedene Typen von Listen:

- Geordnete Listen, die Listeneinträge werden automatisch numeriert
- Ungeordnete Listen, markiert oder unmarkiert, ein- oder mehrspaltig
- Beschreibungslisten, die einem Begriff eine Beschreibung zuordnen

2.2.6.1 Geordnete Listen

Durchnumerierte Aneinanderreihungen schließt das Element <OL> (ordered list) ein, die einzelnen Listenelemente beginnen mit <LI> (list item):

```
<OL>
      <LI>Aufwachen
      <LI>Anziehen
      <LI>Frühstücken
      <LI>Zur Arbeit fahren
</OL>
```

ergibt beispielsweise:

1. Aufwachen
2. Anziehen
3. Frühstücken
4. Zur Arbeit fahren

2.2.6.2 Ungeordnete Listen

Bei ungeordneten Listen ersetzt man das <OL> durch das Element <UL> (unordered list):

```
<UL>
      <LI>Aufwachen
      <LI>Anziehen
      <LI>Frühstücken
      <LI>Zur Arbeit fahren
</UL>
```

ergibt beispielsweise:

- Aufwachen
- Anziehen
- Frühstücken
- Zur Arbeit fahren

2.2.6.3 Beschreibungslisten

Beschreibungslisten erfordern das paarweise Auftreten eines Bezeichners (Term) und seiner Definition. Die Beschreibungsliste wird von einem <DL> Element umschlossen. Die Elemente in einer Beschreibungsliste sind:

DT Der Term
DD Die Definition

Das Beispiel nutzt eine Beschreibungsliste in einem Textbuchausschnitt eines Schauspiels:

```
<DL>
      <DT>King Henry
      <DD>I myself heard the King say he would not be ransomed.
      <DT>Williams
      <DD>Ay, he said so, to make us fight cheerfully: but when our
      throats are cut he may be ransomed, and we none the wiser.
      <DT>King Henry
      <DD>If I live to see it, I will never trust his word after.
      <DT>Williams
      <DD>You pay him then! That's a perilous shot out of an
      elder-gun, that a poor and a private displeasure can do
      against a monarch! You may as well go about to turn the sun
      to ice, with fanning in his face with a peacock's feather.
      You'll never trust his word after! Come `tis a foolish
      saying.
</DL>
```

Die Darstellung könnte wie folgt aussehen:

King Henry
: I myself heard the King say he would not be ransomed.

Williams
: Ay, he said so, to make us fight cheerfully: but when our throats are cut he may be ransomed, and we none the wiser.

King Henry
: If I live to see it, I will never trust his word after.

Williams
: You pay him then! That's a perilous shot out of an elder-gun, that a poor and private displeasure can do against a monarch! You may as well go about to turn the sun to ice, with fanning his face with a peacock's feather. You'll never trust his word after! Come `tis a foolish saying.

2.2.7 Anker

Die Anker (Anchors) sind die Ausgangs- oder Endpunkte eines Hypertextlinks. Der Text zwischen dem Startzeichen und dem Endezeichen ist der Ausgangs- oder Endpunkt eines Links. Zwei wichtige optionale Attribute des Ankerzeichens seien hier kurz erläutert:

HREF Bei Auftreten des HREF Attributes wird der Anker zu sensitivem Text, dem Startpunkt einer Verzweigung. Wenn der Leser des Dokuments diesen Text auswählt, soll ihm das Dokument präsentiert werden, dessen Netzwerkadresse in dem Wert des HREF Attributes definiert wird. Dieses

Dokument muß kein HTML Dokument sein, es kann auch eine Grafik oder ein Soundfile sein, das anschließend in einem externen Betrachter bearbeitet werden kann, oder kann eine Verzweigung zu einem Gopherserver, Newsserver u.a.m. sein. Das Format der Netzwerkadresse wird in dem Abschnitt ‘Universal Resource Locator (URL)’ auf Seite 17 beschrieben. Ein HREF Attribut der Form ‘HREF="#identifier"‘ ermöglicht es, zu dem mit ‘identifier’ benannten Anker in dem aktuellen Dokument zu springen.

NAME Das NAME Attribut erklärt den Anker zu dem Endpunkt einer Verzweigung. Der Wert des Attributes dient gleichzeitig der Identifizierung des Ankers. Der Wert des Attributes ist ein String und darf pro Dokument nur einmal vergeben werden.

Beispiele:

```
See <A HREF="http://info.cern.ch/">CERN</A>'s information for more details.

Information about <a
href="http://www.ncsa.uiuc.edu/demoweb/demo.html#exper"> Experimen-
tal/Advanced Hypermedia Demos </a> are available in the Hypertext demo
document at the NCSA.

A <A NAME="serious">serious</A> crime is one which is associated with im-
prisonment.

                        …

The Organization may refuse employment to anyone convicted of a <a
href="#serious">serious</A> crime.
```

2.2.8 Einbinden von Grafiken und Bildern (IMG)

Status: Extra.

Das IMG Attribute (image) erlaubt es, ein anders Dokument in das aktuelle Dokument einzufügen. Das ist normalerweise eine Grafik oder ein Icon, **nicht** erlaubt ist das Einbetten eines anderen HTML-Dokuments mittels des IMG Elementes. Eine Grafik muß in den Formaten Compuserve GIF oder X-Bitmap (XBM, XPM) vorliegen, andere Formate, wie das hochkomprimierte JPEG Format, werden (zur Zeit) von den Browsern nicht unterstützt[6].

Browser, die keine Inline-Grafiken anzeigen können, ignorieren das IMG Element. Das IMG Element besteht nur aus dem Startzeichen, es hat weder Inhalt noch ein Endezeichen. Es besitzt drei Attribute:

[6] Aufgrund geänderter Lizenzbestimmungen für das GIF-Format ist eine Änderung des verwendeten Grafikformates zu erwarten. Viele (neue) WWW-Browser unterstützen schon das JPEG-Format (25. Mai 1995)

SRC Der Wert des Attributes ist der URL des Dokuments, das eingebunden werden soll. Die Syntax ist entspricht dem HREF Attribut das <A> Elementes.

ALIGN Die Werte des Attributes (TOP, MIDDLE oder BOTTOM) definieren, ob der obere Rand, die Mitte oder der unter Rand der Grafik für die vertikale Ausrichtung verwendet werden.

ALT Alternativer Text, der anstelle der Grafik auf nicht graphischen Systemen angezeigt wird.

Beispiel:

```
Warning: <img src="triange.gif" alt="Warning:"> This must be done by a
qualified technician.

<a href="Go"><img src="Button.gif"> Press to start</a>
```

2.2.9 Erstellen von Tabellen (Table)

Status: Extra.

Ab der Mosaic-Version 2.5 (X11), WinMosaic 2.0.0a8 sowie Netscape 1.1 ist die Darstellung von Tabellen mittels des Table-Statements möglich. Die Definition einer Tabelle wird mit der Anweisung <TABLE>eingeleitet. Soll die Tabelle mit den Begrenzungslinien zwischen den einzelnen Feldern dargestellt werden, ist das Attribut BORDER dem Statement <TABLE> hinzuzufügen.

Soll die Tabelle eine Überschrift oder einen Titel erhalten, ist dies in der Anweisung <CAPTION>…</CAPTION> anzugeben. Der Text in dieser Anweisung wird üblicherweise zentriert und in der Schriftart wie der normale Text angezeigt. Soll die Beschriftung unterhalb der Tabelle erfolgen, ist dem einleitenden <CAPTION>-Element das Attribut ALIGN=BOTTOM zuzuordnen. Innerhalb des <CAPTION>-Elements sind weitere HTML-Elemente erlaubt, z.B. zum Einfügen von Bildern, Schrifthervorhebungen und Anker.

In der HTML wird eine Tabelle zeilenweise definiert. Die Definition einer Zeile wird mit dem Statement <TR> …</TR> (Table Row) umschlossen. Beim <TR>-Element können auch Attribute für die Ausrichtung der ganzen Zeile Angegeben werden. Diese Attribute sind:

- ALIGN=(LEFT, CENTER, RIGHT) zur horizontalen Ausrichtung und
- VALIGN=(TOP, MIDDLE, BOTTOM) zur vertikalen Ausrichtung.

Standardmäßig ist der Text vertikal zentriert angeordnet. Diese Attribute können aber durch die Angabe von Attributen in den einzelnen Zellen überschrieben werden.

Innerhalb einer Zeile werden dann die einzelnen Zellen definiert. Es werden zwei Arten von Zellen unterschieden:

- die Überschriften-Zelle <TH>…</TH> (**T**able **H**eader) und
- die Daten-Zelle <TD>…</TD> (**T**able **D**ata).

Text in der Überschriften-Zelle wird fett und zentriert dargestellt, der Text in einer Datenzelle wird linksbündig und mit normaler Schrift angezeigt. In beiden Zellen wird er vertikal zentriert angeordnet.

Innerhalb der Zellen können weitere HTML-Elemente erscheinen, z.B. zur Texthervorhebung, zum Einfügen von Bildern und Anker u.ä.

Beim <TH> bzw. <TD> Element können auch weitere Attribute zur Ausrichtung des Textes innerhalb der Zelle eingefügt werden. Diese Attribute überschreiben dann die Attribute, die bei der Zeilendefinition im <TR>-Element angegeben wurden. Zur Textausrichtung sind die Attribute ALIGN=(LEFT, CENTER, RIGHT) und VALIGN=(TOP, MIDDLE, BOTTOM) vorgesehen.

Soll der Text innerhalb einer Zelle nicht umgebrochen werden, ist das Attribut NOWRAP anzugeben.

Wie breit eine Tabellenspalte angezeigt wird, entscheidet das Anzeigeprogramm.

Eine Zelle kann auch über mehrere Zeilen oder Spalten der Tabelle gehen. Mit dem Attribut COLSPAN=X wird angegeben, wieviele Spalten die Zelle breit sein soll. Das Attribut ROWSPAN=X gibt an, wieviele Zeilen hoch eine Zelle sein soll.

Ein Beispiel für eine Tabelle ist:

```
<TABLE BORDER>
<CAPTION ALIGN=BOTTOM><A HREF="#tab_verzeichnis">Tab. 1: Beispiel
</A></CAPTION>

<TR>
<TH>Artikel Nr.</TH>
<TH>Beschreibung</TH>
<TH>Menge</TH>
<TH>Datum</TH>
</TR>

<TR>
<TD ALIGN=CENTER>1</TD>
<TD><img src="tomato.gif">Tomaten</TD>
<TH>15</TH>
<TD ROWSPAN=4>15. Jan. 95</TD>
</TR>

<TR>
<TD ALIGN=CENTER></TD>
<TD COLSPAN=2>Nicht verfügbar</TD>
</TR>

...

</TABLE>
```

Auf diesem WWW-Server sind schon eine Reihe von Tabellen definiert worden, so das weitere Beispiele hier nicht aufgeführt werden.

2.3 Universal Resource Locator (URL)

Zur Adressierung von Internet-Ressourcen werden im WWW Universal Resource Locator (URL) verwendet. Einige Ressourcen wurden bereits im Kapitel 1 detailliert beschreiben, weiter sollten noch Daten auf einem FTP-Server oder eine Telnet-Sitzung auf einem entfernten Rechner genannt werden.

Im Zuge der Entwicklung des WWW hatte WWW-Vater Tim Berners-Lee die Idee, diese verschiedenen Objekte mit Hilfe einer einheitlichen Syntax zu adressieren, um sie in Hypertextlinks angeben zu können - den URL's. Sie setzen sich aus einem allgemeinen und einem spezifischen Teil zusammen. Ersterer beschreibt den Bereich bzw. das Namensschema des zu adressierenden Objektes (HTTP, FTP, WAIS, Gopher und so weiter) und der spezielle Teil bezeichnet das Objekt in einer bereichsspezifischen Syntax.

Objekte im Internet verlangen im Anschluß an den allgemeinen Teil mit der Protokollspezifikation die Angabe des Hosts als Domain- oder IP-Adresse und eine Port-Nummer, sofern sie von der für das jeweilige Protokoll vordefinierten Portnummer abweicht. Der Zugriff auf manche Objekte benötigt die Angabe von Benutzerkennung und Paßwort.

Die Kodierung von Sonderzeichen in URL's erfolgt in der Form '%xx', wobei 'xx' für den ISO-8859-1-Code des Zeichens steht, also zum Beispiel '%20' für das Leerzeichen und '%26' für das Prozentzeichen.

Hier eine Übersicht über die zur Zeit definierten Namensschemata:

http: Hypertextdokumente auf WWW-Servern. Beispiele:

```
http://www.informatik.uni-rostock.de/
http://www.mit.edu:8001/usma.html
```

ftp: Die Datei liegt auf einem FTP-Server. Beispiel:

```
ftp://ftp.informatik.uni-rostock.de/pub/magazine/ct/ct94/ct9405.zip
```

news: Der WWW-Client bezieht die gewünschten Artikel von dem in der Umgebungsvariable NNTPSERVER angegebenen Server per NNTP. Beispiele:

`news:comp.infosystems.www`	(Alle Artikel der Newsgruppe)
`news:9311120229.AA24657@expo.x.org`	(Artikel)
`news:*`	(Liste aller Newsgruppen auf den Server)

wais: Anfrage an eine WAIS-Datenbank. Beispiel:

```
wais://gopher.uwo.ca/world-factbook93?Germany
```

Diese Anweisung sucht nach dem Stichwort 'Germany' im CIA World Factbook in der WAIS-Datenbank world-factbook93 auf den Host gopher.uwo.ca.

gopher: Zugriff auf Gopher-Menüs oder Dokumente. Beispiel:

```
gopher://ceylon.informatik.uni-rostock.de/11/Fachbereich%20Informatik
```

telnet: Nutzen eines per Telnet angebotenen Dienstes. Beispiel für den Benutzer ‘user1’ mit dem Paßwort ‘secret’:

```
telnet://user1:secret/trash.informatik.uni-rostock.de
```

Achtung: Die Mosaic-Implementierung ist hier fehlerhaft, da das Programm über den Doppelpunkt zwischen Usernamen und Paßwort stolpert.

Es existieren noch weitere Namensschemata und die Liste ist außerdem offen für weitere neue Dienste.

2.4 Style Guide

Im Style Guide soll dem HTML Autor eine (kurze) Anleitung zur Erstellung gut lesbarer Dokumente gegeben werden.

Informationen sollten gut strukturiert werden, normalerweise ist das ein hierarchischer Baum mit einzelnen Kapiteln wie in einem Buch. Die einzelnen Dokumente sollten aber nicht zu groß sein. Einerseits, weil lange Dokumente auch eine lange Zeit zur Übertragung benötigen und andererseits kann dem Leser das Scrollen durch das Dokument schwer fallen, z.B. an ASCII-Terminals. Online Hilfstexte und Indexseiten, die im wesentlichen nur in andere Dokumente verzweigen, sollten nicht länger als 24 Zeilen sein. Dieses überprüft man am besten an einem ASCII-Terminal. Reine Textdokumente sollten maximal 5 A4 Seiten füllen.

Jedes Dokument sollte vom Autor unterschrieben werden, z.B. mit einer Verzweigung zu einer Seite, auf der der Autor sich vorstellt. Außerdem sollte der Status eines Dokuments angeben werden, z.B. wann das Dokument ungültig wird oder in welchem Zustand der Bearbeitung es sich befindet. Die Angabe des Erstellungsdatums und des Zeitpunktes der letzten Modifikation ist ebenfalls für den Leser wichtig.

Der Hauptunterschied zwischen dem Schreiben eines (gedruckten) Textes, z.B. eines Buches, und einer Serie von (zusammengehörigen) Online-Dokumenten ist, daß der Leser in das Dokument von irgendwo hineingesprungen sein kann. Die Überschrift und der Dokumententitel sollten dem Leser immer einen Überblick über das jeweilige Dokument geben. Desweiteren sollte ein Dokument niemals mit Formulierungen wie „Das nächste Thema was wir betrachten ist…“ oder „Die einzige Lösung zu dem Problem ist…“ beginnen; sie würden den Leser nur verwirren. Manchmal können Einleitungsworte mit Background-Informationen in einem anderen Dokument verbunden werden. Ebenso ist es wichtig sich vorzustellen, was man selbst aus dem Dokument erfahren will, wenn man es nutzt.

HTML-Dokumente enthalten keine Informationen über die Schriftart und -type in der das Dokument angezeigt werden soll. Das hat den großen Vorteil, daß es auf allen Plattformen erfolgreich angezeigt werden kann, auch am ASCII-Terminal. Die Dokumente sollten deshalb auch auf allen Plattformen mit möglichst vielen Browsern betrachtet werden. Es nützt nichts, ein Dokument speziell auf einen Browser abzustimmen, z.B. mit extra Leerzeichen und -zeilen. Aufgrund der Geräteunabhängigkeit sollten auch Formulierungen wie „Klicken Sie hier“ vermieden werden, der Leser am ASCII-Terminal besitzt nämlich keine Maus.

Jedes Hypertextdokument sollte wie ein Programm getestet werden. Wenigstens einer aus der Gruppe der Interessenten des Dokuments sollte es lesen und seine Eindrücke wiedergeben. Darum sollte man auch angeben, wie der Leser seine Kommentare und Kritiken an den Autor senden kann.

Im Logfile des WWW-Servers kann der Autor etwas über die Leser seines Dokuments erfahren und ob seine Dokumente überhaupt gelesen werden. Wenn Dokumente nicht gelesen werden, kann es daran liegen, daß die Leser sie nicht finden können, weil sie auf den Text nicht aufmerksam gemacht werden.

2.5 Erweiterungen von HTML zu HTML+[7]

Drei Jahre praktische Erfahrungen mit der HTML haben eine Reihe von Schwachpunkten bloßgelegt. Autoren möchten häufig mehr Einfluß auf die Darstellung des Werkes nehmen können, abweichend von der SGML-Lehre, Form und Inhalt streng voneinander zu trennen. Ein simpler Zeilenvorschub ist beispielsweise nicht möglich, ohne gleich einen neuen Absatz zu beginnen. Text in einem PRE Element berücksichtigt zwar den Zeilenumbruch, aber der Text erscheint in einem nichtproportionalen Font wie Courier, was dann häufig noch schlimmer aussieht. Rechtsbündige Texte oder ein Blocksatz lassen sich auch nicht darstellen. Diese Auswahl ist nur willkürlich, besonders schmerzlich ist aber, daß keine Tabellen definierbar sind und sie deshalb in häßlichen PRE-Blöcken angegeben werden müssen. Um diese und andere Schwachpunkte zu umgehen, wird gegenwärtig im Internet über einen Standard für den HTML Nachfolger HTML+ diskutiert.

Vieles davon befindet sich noch im Stadium der Diskussion und wurde noch nicht in einem WWW-Client implementiert. Einige Features, wie interaktive Formulare, die *Mosaic 2.x* unterstützt, kann man bereits schon heute nutzen. Der Zeilenvorschub geschieht in HTML+ mit dem Tag <BR>, eine horizontale Linie als optische Trennung erreicht der Autor durch <HR>. Auch diese Features sind bereits in *Mosaic 2.x* implementiert.

HTML+ soll auch die Möglichkeit zur Darstellung von Tabellen mittels des <TABLE> Elementes und der Gestaltung mathematischer Formeln innerhalb des <MATH> Elementes bieten. Es sollen auch mehr Möglichkeiten zur Textgestaltung, wie rechtsbündiger Text und Blocksatz möglich sein. Auch die Menge der verfügbaren Zeichen wird erweitert und nicht mehr auf ISO 8859-1 beschränkt sein.

3 HTML-Beispiele

Einige einfache Beispiele zu diesem Text befinden sich im Home-Verzeichnis des Autors auf den Rechnern des Fachbereichs Informatik der Universität Rostock, z.B. im Verzeichnis ~garlipp/html[8]. Da die meisten WWW-Browser das Betrachten und Abspeichern von HTML-Dokumenten im Quellformat ermöglichen, ist es leicht, sich auch andere interessante Dokumente im WWW als HTML-Quelle zu betrachten.

---

[7]Die Bezeichnung HTML+ wird zukünftig nicht mehr verwendet. Der gegenwärtige De-facto-Standard soll als HTML 2.0 normiert werden. HTML 3.0 soll als wesentliche Neuerung Tabellen unterstützen, mit HTML 3.1 sollen auch mathematische Formeln definiert werden. Ein erster HTML 3.0-Browser wird für Herbst 1994 erwartet. (24. Juni 1994)

[8]Der Autor betreut z.Z. zwei WWW-Server auf *www.informatik.uni-rostock.de* und *www-dbis.informatik.uni-rostock:80* (WWW-Server des Lehrstuhls Datenbanken und Informationssysteme). (25. Mai 1995)

4 Literatur

Einige Textabschnitte und Beispiele wurden den folgenden Quellen entnommen:

4.1 NetNews

- Wolfgang Sander-Beuermann; Usenet; Lese- und Schreibgewohnheiten der Benutzer; iX 10/93, S. 56 ff.

4.2 Gopher

- Dokumente des Gopherservers der University of Minnesota (gopher.tc.umn.edu)
- Ingo Dreßler; Information Retrieval, Das Internet-Werkzeug Gopher; iX 4/93, S. 56 ff.

4.3 WAIS

- Erich Weichselgartner; WWW: Hypertext-basiertes Informationswerkzeug im Internet; iX 12/93, S. 84 ff.
- Nathan Torkington (gnat@kauri.vuw.ac.nz); WAIS Introduction; Artikel aus der Newsgruppe comp.archives.admin; gopher*://saturn.soils.umn.edu:70/0R0-5407-/wais.what-is*

4.4 WWW und HTML

- Erich Weichselgartner; WWW: Hypertext-basiertes Informationswerkzeug im Internet; iX 12/93, S. 84 ff.
- Rainer Klute; WWW-Frontend; Internet-Informationsdienste mit Mosaic; iX 2/94, S. 150 ff.
- Rainer Klute; Hypertext; HTML und HTTP; iX 9/94, S. 170 ff.
- Rainer Klute; Serviertechnik; Installation und Konfiguration eines CERN-WWW-Servers; iX 3/94, S. 182
- Tim Berners-Lee, Daniel Connolly; Hypertext Markup Language; A Representation of Textual Information and Metainformation for Retrival and Interchange; Internet Draft, 13 Juli 1993; *ftp://ftp.w3.org/pub/www/doc/html-spec.ps*
- Tim Berners-Lee; Hypertext Transfer Protocol; A Stateless Search, Retrive and Manipulation Protocol; Internet Draft; 5. November 1993; *ftp://info.cern.ch/pub/www/doc/http-spec.ps*
- Tim Berners-Lee; World-Wide Web server software; Generated from the Hypertext; 2. Juni 1993; *ftp://ftp.w3.org/pub/www/doc/www-server-guide.ps*
- David Raggett; HTML+ (Hypertext Markup format) - Internet Draft; *ftp://hplose.hpl.hp.com/pub/draft-ragget-www-html-00.ps*
- Tim Berners-Lee; Uniform Resorce Locatores; A Unifying Syntax for the Expression of Names and Addresses of Objects on the Network; Internet Draft; 14. Oktober 1993; *ftp://info.cern.ch/pub/www/doc/url7a.ps*
- Rainer Klute; Erste internationale WWW-Konferenz in Genf; iX 7/94, S. 14f. (24. Juni 1994)