

Kurzbeschreibung der Sprache HTML

HTML (HyperText Markup Language) ist eine Sprache zur Beschreibung der Struktur von Hypertext-Dokumenten. Sie ist definiert als eine Applikation der Standard Generalized Markup Language (SGML ISO Standard 8879). Seit 1990 wird diese Sprache für die Dokumente im globalen Informationssystem World Wide Web genutzt. Im Laufe der Zeit hat es eine Entwicklung der Spezifikation von HTML 1.0 über HTML 2.0 zu HTML 3.0 gegeben. Diese letzte Spezifikation ist zur Zeit noch in der Phase der Definition. In dieser Kurzbeschreibung sollen im wesentlichen die Elemente von HTML 2.0 erläutert werden. Der Autor von HTML-Dokumenten sollte berücksichtigen, daß diese Sprache die Struktur des Textes beschreibt. Das Erscheinungsbild des Dokumentes wird zum Teil vom Benutzer (Leser) und von den Leseprogrammen (Browsern) bestimmt.

Die HTML-Dokumente sind normale Texte, die aber zusätzlich spezielle Zeichenketten, die HTML-Befehle (Tags), enthalten. Die Befehle beginnen immer mit dem Zeichen "<", dann folgt der Name des Befehls und danach das Zeichen ">" (z.B. <H1>). Zusätzlich dazu können Befehle noch Attribute enthalten (z.B.). Einige HTML-Elemente bestehen aus einem Start- und einem Endebefehl. Der Endebefehl wird durch Voranstellen des Zeichens "/" vor dem Namen gekennzeichnet (z.B. </H1> ist das Ende von <H1>).

Die allgemeine Struktur eines Dokumentes sollte folgende Form haben:

```
<HTML>
<HEAD>
<TITLE>Titel des Dokumentes</TITLE>
</HEAD>
<BODY>
Text des Dokumentes
</BODY>
</HTML>
```

Die Befehle <HTML> und </HTML> kennzeichnen das Dokument als HTML-Text. Es folgt der Kopf des Dokumentes, eingeschlossen durch die Befehle <HEAD> und </HEAD>. In diesem Bereich befinden sich nur Informationen über das Dokument (z.B. der Titel des Dokumentes). Der eigentliche Inhalt wird eingeschlossen durch die Befehle <BODY> und </BODY>.

Befehle im Kopf des Dokumentes

Titel

Der Befehl <TITLE> kennzeichnet den Titel des Dokumentes. Der Titel wird häufig für Listen (Hotlist, History) oder Suchfunktionen benutzt. Deshalb

ist es notwendig, den Titel so zu wählen, daß er eine Aussage zum Inhalt des Dokumentes enthält.

Beispiel:

```
<TITLE>Beschreibung von HTML</TITLE>
```

Isindex-Suche

Der Befehl <ISINDEX> legt fest, daß das Dokument vom Leser zur Index-Suche genutzt werden kann. Im Dokument erscheint dadurch ein Feld, das den Leser zur Eingabe einer Phrase auffordert. Voraussetzung für die Suche ist, daß der Server ein Suchprogramm zur Unterstützung dieses Befehls besitzt.

Beispiel:

```
<ISINDEX>
```

Informationen über das Dokument

Das Element <META> wird genutzt, um zusätzliche Informationen über das Dokument bereitzustellen. Diese Information kann für die Katalogisierung, Schlagwortsuche u.a. eingesetzt werden.

Beispiel:

```
<META NAME="Keywords" CONTENT="HTML
Editor WWW">
```

Befehle im Bereich <BODY>

Überschriften

Es gibt sechs Stufen von Überschriften, dargestellt durch die Befehle <Hn> (n=1..6). Diese Überschriften dienen nur zur Strukturierung des Textes, d.h. der Autor legt nur die Stufe der Überschriften fest, nicht aber deren Aussehen.

Beispiel:

```
<H1>HTML - Sprache des World Wide Web</H1>
```

Absatz und Umbruch

In HTML-Dokumenten werden Tabulatoren, mehrere Leerzeichen oder Zeilenumbrüche immer als ein Leerzeichen interpretiert. Bei der Darstellung des Dokumentes wird durch das Leseprogramm ein automatischer Zeilenumbruch am rechten Fenster- oder Bildschirmrand durchgeführt. Zur Strukturierung des Textes werden deshalb spezielle Befehle von HTML eingesetzt.

Der Befehl <P> trennt zwei Absätze voneinander. An dieser Stelle wird ein Zeilenumbruch und im allgemeinen ein Zwischenraum erzeugt.

Beispiel:

```
Hier ist ein Absatz zu Ende.<P>
```

Es beginnt ein neuer Absatz.

(Das gleiche Ergebnis hat folgender Ausdruck:
Hier ist ein Absatz zu Ende.<P>Es beginnt ein neuer Absatz.)

Durch den Befehl `
` wird ein Zeilenumbruch erreicht.

Beispiel:

Hier ist eine Zeile.`
`

Jetzt beginnt eine neue Zeile.

Wenn mehrere Befehle zum Zeilenumbruch oder Absatz direkt aufeinander folgen, wird im allgemeinen nur einer berücksichtigt. Deshalb ist dies nicht geeignet, um einen größeren vertikalen Abstand im Dokument zu erzeugen. Es ist besser, die vertikale Strukturierung durch den Befehl `<HR>`, der eine horizontale Linie im Dokument erzeugt, vorzunehmen.

Beispiel:

Hier ist das Ende eines Abschnittes

`<HR>`

Dies ist ein neuer Abschnitt.

Texte, die eine feste Formatierung in Zeilen und Spalten erfordern, können mit dem Befehl `<PRE>` als vorformatiert (preformatted) gekennzeichnet werden. Hier bleiben dann alle Leerzeichen und Zeilenumbrüche des ursprünglichen Textes erhalten. Bei der Darstellung des Textes wird kein automatischer Zeilenumbruch durchgeführt. Dieser Text wird im allgemeinen mit einem nichtproportionalen Zeichensatz dargestellt. Mit diesem Befehl können auch tabellenartige Texte gezeigt werden.

Beispiel:

`<PRE>`

```
Name      |Telefon
-----+-----
Maier     |203 38 88
Kohl      |210 09 56
```

`</PRE>`

Eine besondere Art des Abschnittes ist ein Zitat oder Beispiel. Gekennzeichnet wird dieser Abschnitt durch den Befehl `<BLOCKQUOTE>`. Häufig wird der Text eingerückt und/oder kursiv angezeigt.

Beispiel:

`<BLOCKQUOTE>`

Das ist ein Zitat.

`</BLOCKQUOTE>`

Listen

HTML bietet verschiedene Möglichkeiten, Informationen in Form von Listen darzustellen.

Die geordnete Liste (Ordered List) präsentiert eine Anzahl von Elementen, die in einer Reihenfolge geordnet sind. Sie wird durch die Befehle `` und `` eingerahmt. Die Listenelemente werden durch `` gekennzeichnet.

Beispiel:

Die Installation von Mosaic wird folgendermaßen durchgeführt:

``

`Installieren von Win32s`

`Installieren von Mosaic`

`Installieren der Viewer`

``

Die Installation von Mosaic wird folgendermaßen durchgeführt:

1. Installieren von Win32s
2. Installieren von Mosaic
3. Installieren der Viewer

Wenn die Elemente einer Liste keiner Reihenfolge unterliegen, kann die ungeordnete Liste (Unordered List) genutzt werden. Sie unterscheidet sich in der Syntax von der vorhergehenden nur durch die Rahmenbefehle `` und ``. In der Darstellung werden statt der Nummern vor den Listenelementen neutrale Zeichen (Punkte, Striche) gezeigt.

Beispiel:

Als Browser werden genutzt:

``

`Mosaic`

`Netscape`

`WinWeb`

``

Als Browser werden genutzt:

- Mosaic
- Netscape
- WinWeb

Ein besonderer Typ von Listen ist die Definitionsliste. Sie besteht aus einer Anzahl von Schlüsselwörtern mit ihren entsprechenden Definitionen. Diese Liste wird durch die Befehle `<DL>` und `</DL>` eingerahmt. Die Schlüsselwörter werden durch `<DT>` gekennzeichnet. Nach diesen Schlüsselwörtern folgt die entsprechende Erläuterung, die mit dem Befehl `<DD>` beginnt.

Beispiel:

`<DL>`

`<DT>Begriff 1`

`<DD>Erläuterung zum Begriff 1`

`<DT>Begriff 2`

`<DD>Erläuterung zum Begriff 2`

`</DL>`

Begriff 1

Erläuterung zum Begriff 1

Begriff 2

Erläuterung zum Begriff 2

Befehle zur Zeichenformatierung

Zur Hervorhebung von Textstellen können Befehle zur Änderung des logischen oder physischen Erscheinungsbildes der Zeichen genutzt werden. Sie bestehen aus einem Start- und einem Endebefehl.

Zur Änderung der logischen Erscheinung eines Textes werden die Befehle *Emphasis* (Akzentuierung) `` und `` und *Betonung* `` und `` genutzt.

Beispiel:

```
<EM>hervorgehobener Text</EM>
<STRONG>betonter Text</STRONG>
```

Die Befehle zur logischen Struktur bewirken eine Hervorhebung des Textes, wobei die Art der Hervorhebung (fett, kursiv, farbig) vom Browser oder von dessen Konfiguration abhängt. Im allgemeinen wird aber der Text bei `` kursiv und bei `` fett dargestellt.

Im Gegensatz dazu schreiben die Befehle zur Änderung des physischen Aussehens des Textes das Erscheinungsbild der Buchstaben vor.

Beispiele:

```
<I>italic kursiv</I> für kursiv
<B>Fettdruck</B> für Fettdruck
```

Anker

Zur Strukturierung der Information bietet die Sprache HTML die Möglichkeit, in den Dokumenten Querverweise auf andere Objekte (Text, Bilder, Ton- und Video-Sequenzen) zu integrieren. Diese Querverweise werden als Hypertext-Verbindungen bezeichnet. Mit dem Anker-Befehl `<A>` und `` werden Textpassagen gekennzeichnet, die als Start oder Ziel dieser Hypertext-Verbindungen dienen. Enthält dieser Befehl das Attribut `HREF`, so wird der Anker als Verweis auf ein Dokument, das durch die URL [1] [2] des Attributes bestimmt wird, interpretiert.

Beispiel:

```
<A HREF="http://www.rz.hu-berlin.de/index.html">Home
Page der HUB</A>
```

In diesem Beispiel wird der Text "Home Page der HUB" vom Browser hervorgehoben. Bei der Auswahl dieses Textes durch den Leser wird eine Verbindung zum Rechner (Server) "www.rz.hu-berlin.de" hergestellt und das Dokument "index.html" gelesen.

Durch das Attribut `NAME` wird der Anker zu einem Ziel für Hypertext-Verbindungen innerhalb eines Dokumentes.

Beispiel:

```
Hier sind Informationen zur <A NAME="geschichte">Geschichte</A>
```

Ein anderes Dokument kann auf diesen Anker verweisen, indem der Name an die URL angehängt wird.

Der Befehl: `Geschichte` verweist z.B. auf den obengenannten Anker.

Bilder

Mit dem Befehl `` können Bilder in ein Dokument eingebunden werden. Der Wert des Attributes `SRC` stellt die URL der Bilddatei dar. Das Attribut `ALIGN` mit den Werten `TOP`, `MIDDLE` oder `BOTTOM` bestimmt die Position des nachfolgenden Textes. Wenn der Wert nicht angegeben ist, wird der Text am unteren Rand des Bildes ausgerichtet. Optional kann das Attribut `ALT` einen Text mit dem Titel des Bildes oder einer Erläuterung enthalten. Dieser Text erscheint dann, wenn das Bild nicht dargestellt werden kann.

Beispiel:

```
<IMG SRC="rz.gif" ALT="Logo des Rechenzentrums">
```

Sonderzeichen

Die Definition der Sprache HTML verlangt, daß die Dokumente nur aus Zeichen des (7-Bit) ASCII-Zeichensatzes bestehen. Dadurch wird gewährleistet, daß diese Dokumente ohne Probleme zwischen verschiedenen Computern ausgetauscht werden können. Da dieser Zeichensatz nur die Zeichen des englischen Alphabets, die Ziffern und einige Sonderzeichen enthält, müssen darüber hinausgehende Zeichen (z.B. deutsche Umlaute) durch sogenannte Entities umschrieben werden. Entities beginnen immer mit dem Ampersand (&), dann folgt die Umschreibung und zum Abschluß ein Semikolon (;).

Beispiele:

<code>&auml;</code> für ä,	<code>&Auml;</code> für Ä
<code>&ouml;</code> für ö	<code>&Ouml;</code> für Ö
<code>&uuml;</code> für ü	<code>&Uuml;</code> für Ü
<code>&szlig;</code> für ß	<code>&amp;</code> für &
<code>&lt;</code> für <	<code>&gt;</code> für >

Eine Gesamtliste der Entities ist in [3] zu finden.

Dieser Artikel zeigt nur die wichtigsten Befehle der Sprache HTML 2.0. Eine ausführliche Beschreibung ist im Internet-Draft [4] zu finden.

In einer Weiterentwicklung der Sprache (HTML 3.0) wird es dann auch möglich sein, Tabellen, mathematische Gleichungen und Formeln darzustellen.

Lothar Wendroth

Literatur:

- [1] Naumann, W.: World Wide Web. RZ-Mitteilungen, Nr. 8, Oktober 1994.
- [2] Was ist ein URL? <http://www.informatik.hu-berlin.de/Rechnerbetrieb/glossar/url.html>
- [3] iso8859-1 table, <http://www.informatik.hu-berlin.de/Themen/www/ISOlat1.html>
- [4] Berners-Lee, T., und Connolly, D.: `<draft-ietf-html-spec-02.txt>`, Mai 95
<http://www.w3.org/hypertext/WWW/MarkUp/html-spec/html-spec.txt>

Anhang:

Das abschließende Beispiel soll die Struktur eines Dokuments und die Wirkung einiger Befehle illustrieren.

```
<HTML>
<HEAD>
<TITLE>Beispiel-Dokument</TITLE>
<META NAME="AUTHOR" CONTENT="L. Wendroth">
</HEAD>
<BODY>
<H1><IMG ALIGN=BOTTOM SRC="achtung.gif" ALT="Achtung !!"> Beispiel</H1>
Ein <EM>kleines</EM> Beispiel soll die <STRONG>Wirkung</STRONG> einiger HTML-Befehle zeigen.<BR>
In diesem Dokument sind enthalten:
<UL>
<LI>Listen
<LI>Entities
<LI>Zeichenformatierung
</UL>
Dazu die Erläuterungen:
<DL>
<DT>HTML
<DD>Sprache zur Beschreibung von Hypertext-Dokument
<DT>Browser
<DD>Programm zum Lesen (Betrachten) von HTML-Dokumenten
</DL>
<HR>
<PRE>
Ein vorformatierter Text
    (preformatted Text)
        wird so dargestellt,
            wie er im Original formatiert wurde.
</PRE>
<HR>
Weitere Informationen aus dem Rechenzentrum finden Sie auf unserem <A HREF="http://www.rz.hu-berlin.de/inside/rz/">Server</A>.
</BODY>
</HTML>
```



Beispiel

Ein *kleines* Beispiel soll die **Wirkung** einiger HTML-Befehle zeigen.
In diesem Dokument sind enthalten:

- Listen
- Entities
- Zeichenformatierung

Dazu die Erläuterungen:

HTML
Sprache zur Beschreibung von Hypertext-Dokument

Browser
Programm zum Lesen (Betrachten) von HTML-Dokumenten

```
Ein vorformatierter Text
    (preformatted Text)
        wird so dargestellt,
            wie er im Original formatiert wurde.
```

Weitere Informationen aus dem Rechenzentrum finden Sie auf unserem [Server](#).