

XHTML

1999 wurde vom WWW-Konsortium (kurz W3C), das für die Festlegung der Sprache für Internet-Dokumente zuständig ist beschlossen, HTML nicht mehr weiterzuführen.

HTML-3 war noch eine wüste Mischung aus Dokumentinformationen und Gestaltungseigenschaften (, , <I>, ... standen mitten im Text herum). In HTML-4 wurde durch die Einführung der Style Sheets eine Trennung versucht, jedoch auf Abwärtskompatibilität zu Version 3 zu viel Wert gelegt. Das hatte zur Folge, dass sich weder Browser-Hersteller noch Webdesigner genötigt sahen, von alten schlechten Gewohnheiten Abstand zu nehmen. Allerdings haben die Browserhersteller einen schwierigen Mittelweg zu gehen: Wenn man nur den neuen Standard unterstützt, sind 99% der Internetseiten unlesbar.

Inzwischen waren die Browser wenigstens so modern und Standard-konform, die neuen Elemente zu einem großen Teil korrekt umzusetzen.

Deshalb wurde beschlossen, kein HTML-5 zu definieren, sondern gleich etwas besseres, nämlich XHTML (Extended Hypertext Markup Language. Eine Spezialform der XML). War es Webautoren bisher nur *möglich*, Seiteninhalte von ihrer Darstellungsform zu trennen, werden sie nun *gezwungen*, dies zu tun.

Die Tatsache, dass auch seit 1999 viele Webdesigner und sogar Programme zur automatischen Seitenerstellung (Frontpage, Fusion, NetscapeComposer,...) Code mit nicht mehr erlaubten Auszeichnungen erstellten, soll diesen Wechsel nicht aufhalten. (Würde man darauf Rücksicht nehmen, könnte man niemals wechseln...)

Freundlicherweise werden sich die Browser aber noch eine Zeit lang mit den alten Schreibweisen abgeben. Weiters hat das W3C eine 'Hintertür' eingebaut, mit der man alte Tags auch in XHTML noch verwenden kann – sie sind aber als 'deprecated' gekennzeichnet – das heißt: sie könnten von heute auf morgen nicht mehr erlaubt sein.

Die Trennung zwischen Inhalt und Gestaltung hat oberste Priorität. Dies gilt übrigens inzwischen auch für andere Anwendungen. OpenOffice verwendet bereits seit langem das verwandte XML-Format, sogar Microsoft hat angekündigt, dieses (offene!) Format zu unterstützen. Moderne Programmiersprachen wie Java oder Python haben die Regeln von XML ebenfalls bereits verinnerlicht.

Eine Schattenseite gibt es leider doch: auch 2003 haben alle Browser noch ihre kleinen Macken – nicht alles funktioniert so, wie es sollte. Gelegentlich werden tags nicht ganz richtig interpretiert, manchmal werden sie sogar ignoriert (etwa beherrscht nur Amalthea den mathematischen Formelsatz).

Folgende wichtige Regeln müssen wir beachten:

- XHTML-Dateien müssen mit dem richtigen Kopf beginnen. Dieser besteht aus einem XML-Prolog (er gibt Typ und Zeichensatz an, optional) und einem XHTML-Header. Einige Browser werden durch den XML-Prolog verwirrt. Momentan lässt man ihn also besser noch weg.
- Tags müssen in Kleinbuchstaben geschrieben werden. Ab nun ist HTML case-sensitive.
- Alle tags müssen geschlossen werden. Also etwa <hr></hr>. Falls ein tag wie hier gar keinen Text umgrenzt, darf die Kurzschreibweise <hr /> Verwendung finden.
- Alle Attribut-Werte sind in Anführungszeichen zu setzen. Also href="seite2.html" statt (wie

bisher immer schon falsch, aber toleriert) href=seite2.html

Zum Kopf der XHTML-Datei:

```
<?xml version="1.0" encoding="UTF-8"?>
```

Das ist der optionale Prolog.

Für den XHTML-Header gibt es drei Ausführungen:

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

Das steht für ein echtes XHTML-Dokument

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

Dies beschreibt ein Dokument, das noch alte Tags enthalten darf.

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

Diese Datei beschreibt ein Frameset. Ein <body>-tag ist hier im Unterschied zu bisher nicht mehr erlaubt!

Übrigens war dieser DOCTYPE bereits in HTML-4 verpflichtend. Allerdings hat kein Browser sein Fehlen beanstandet.

Merkspruch:

Früher hat man gerne HTML 'ausprobiert' und war mit seinem Code zufrieden, wenn die Kontrolle im Browser positiv ausgefallen ist. Diese Zeit ist glücklicherweise vorbei. **Wir müssen wissen, wie eine Eigenschaft definiert wird, und setzen dies erst dann um. Der Test in den Browsern ist kein Test für unseren Code, sondern nur mehr für die Kompatibilität der Browser!**

Links ins Internet:

- Der Opera-Browser hat im Kontextmenü einer Seite den Eintrag Frame/Validate Source. Diese Option schickt die Seite an einen Server des W3C, wo der Sourcecode auf Korrektheit überprüft und eine Fehlerliste zurückgeschickt wird. Aber Achtung: korrekter XHTML-Code wird von den heutigen Browsern nicht 'korrekt' dargestellt. Schreibt man Seiten für reale Browser, ist man zu Unkorrektheiten gezwungen
- Der Browser 'Amalthea' wird vom W3C gepflegt. Er ist absolut 'korrekt', kann deswegen jedoch viele Seiten des realen Internets nicht darstellen (da sie unkorrekte Syntax enthalten).
- Unter <http://www.w3.org/TR/REC-CSS2> findet man eine vollständige Spezifikation der Stylesheets.

Abschließend noch ein Beispiel für eine korrekte Seite mit zugehörigem Stylesheet.

```
<?xml version="1.0" encoding="UTF-8"?>
```

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

<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="de" lang="de">
<head>
  <meta http-equiv="Content-type" content="text/html; charset=UTF-8" />
  <link rel="stylesheet" type="text/css" href="myTestStyle.css" />
  <title>Meine Testseite</title>
</head>

<body>
  <h1>Titel</h1>
  <p class="blau">erster Absatz</p>
  <p class="gross">zweiter Absatz</p>
  <p><div id="kursiv">dritter <span id="gelb">Absatz</span></div></p>
  <p><a href="nix.html">ein link!</a></p>
</body>

</html>
```

Und die Datei myTestStyle.css, die noch einige weitere Beispiele enthält.

```
/* Definition fuer Attribute eines tags */
h1 { font-size:40pt;
     font-style:bold;
     font-family:Arial,sans-serif;
     color:#00ff00;
     }

h2, h3
  { color:#ff00ff;
    font-family:Georgia,serif;
    }

body {
  font-family: Arial,sans-serif;
  background-color:#909090
}

/* Klassen nur fuer den tag p gueltig */
p.gross { font-size:250%; }
p.klein { font-size:75%; }

/* universelle Klasse, ueberall einsetzbar */
.blau { color:#0000ff; }

/* Definition fuer gelegentlichen Einsatz mit ID Selektoren */
#kursiv { font-style:italic; }
#gelb { color:#ffff00; }

/* Pseudo-Elemente */
a:link { color:#FF0000; font-weight:bold }
a:active { color:#0000FF; font-style:italic }
a:visited { color:#990000; }

/* geschachtelte tags: also nur <h2>...<i>HIER</i>...</h2> gueltig*/
h2 i { font-size:90%; }
```