

# Einführung in HTML

## Ein einfaches HTML-Dokument

Die Abkürzung HTML steht für HyperText Markup Language – Auszeichnungssprache für Hypertext. Wir befassen uns hier im wesentlichen mit der Version 4 von HTML.

Ein typisches HTML-Dokument besteht aus Marken (englisch: *tags*) und Text. Die Marken legen fest, ob ein Abschnitt des Dokumentes Titel, Text, Aufzählung oder Überschrift ist.

Jedes HTML-Dokument ist in zwei Abschnitte unterteilt, einen Dokumentenkopf und den Dokumentenkörper. Die dafür verwendeten Marken sind `<head>` und `<body>`.

Im Kopf wird auch der Titel des Dokuments festgelegt. Ein kurzes HTML-Dokument sieht etwa so aus:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN"
  "http://www.w3.org/TR/REC-html40/strict.dtd">
<html>
  <head>
    <title>Mein erstes HTML-Dokument</title>
  </head>
  <body>
    <p>Hallo!</p>
  </body>
</html>
```

Hier stehen die bereits besprochenen Marken `<head>` und `<body>` für Kopf und Inhalt des Dokuments. Beide werden innerhalb des Elements `<html>` zusammengefaßt. Neu sind die Dokumententypklärung ganz am Anfang und die Marken `<title>` und `<p>`. Die Dokumententypklärung gibt an, welche Version des HTML-Standards verwendet wird. Es gibt viele Versionen, bei denen sich auch die Bedeutung der Marken geringfügig unterscheiden kann. Ein Webbrowser kann mit Hilfe der Dokumententypklärung eine für die Darstellung günstigere Interpretation auswählen. Die Marke `<title>` legt den Titel des Dokuments fest. Mit dem Element `<p>` wird ein Paragraph (Abschnitt) markiert.

Die bisher verwendeten Marken gibt es in zwei Versionen; die zweite mit dem vorangesetzten Schrägstrich gibt das Ende des Elementabschnitts an; `</p>` für Paragraphen, `</html>` für das ganze Dokument..

## Einfache Formatierungen – Überschriften

Es gibt sechs verschiedene Überschriftstypen `<h1>`, `<h2>`, `<h3>`, `<h4>`, `<h5>` und `<h6>`. Auch hier gibt es entsprechende Endmarken, mit denen das Ende der jeweiligen Überschrift markiert wird: `</h1>`, `</h2>`, `</h3>`, `</h4>`, `</h5>` und `</h6>`.

Für die Überschriften gibt es weitere Regeln: Die oberste Überschrift ist immer vom Typ *h1*. Die einer Überschrift nachfolgende Überschrift darf höchstens ein Grad unter dem der aktuellen liegen. Das heißt, nach *h3* darf *h4*, aber nicht *h5* oder *h6* kommen. Überschriften höheren Grades sind beliebig erlaubt.

Man kann sich das wie bei einem Buch vorstellen. Ein Buch hat Kapitel, Unterkapitel, Abschnitte und so weiter. Auch hier wird üblicherweise immer nur einen Schritt nach unten gegangen.

Dem Leser mögen die Überschriften zu groß erscheinen und er will lieber kleinere Größen verwenden. Davon sei abgeraten, da der Leser das Aussehen eines Elements selbst bestimmen kann. Macht der Autor eines Dokuments hier feste Annahmen, so wird er den Lesegenuß vieler Leser trüben. In einem späteren Abschnitt werden wir lernen, wie man das Aussehen eines Elements bestimmt und dabei auch die Vorlieben der – möglicherweise von spezieller Software abhängiger – Leser berücksichtigt.

## Aufgaben

Schreibe ein einfaches HTML-Dokument mit verschiedenen Überschriften und Paragraphen.

Teste das Aussehen der Seite mit einem Webbrowser.

## Aufzählungen

Dem gelernten Designer gleich dem interessierten Studenten sind sie ein Grauen: Aufzählungslisten wie sie in unzähligen Vorlesungen Information verwirren und unverständlich machen. In HTML gibt es zwei Typen von Listen, ungeordnete (Bullet-Listen) und geordnete (numerierte Listen, Aufzählungslisten). Ungeordnete Listen werden mit `<ul>` und `</ul>` eingeschlossen, geordnete mit `<ol>` und `</ol>`. Die einzelnen Punkte einer Liste werden wiederum mit `<li>` und `</li>` eingeschlossen, unabhängig davon, ob eine geordnete oder ungeordnete Liste vorliegt. Wir geben ein kurzes Beispiel einer ungeordneten Liste:

```
<ul>
  <li>gelb</li>
  <li>rot</li>
  <li>blau</li>
</ul>
```

## Markierungen für Textblöcke und innerhalb von Texten.

Die bisher eingeführten Markierungen sind – mit Ausnahme von `<li>` – für Textblöcke gedacht. Diese Markierungen dürfen keine anderen Markierungen für Textblöcke enthalten. Ausnahmen sind `<ul>` und `<ol>`; ohne diese Ausnahme sind keine verschachtelten Aufzählungen möglich. Insbesondere sollten Textparagrafen mit `</p>` abgeschlossen werden, bevor eine Aufzählung oder eine Überschrift kommt.

Markierungen innerhalb von Texten unterliegen dieser Einschränkung nicht. Ein Link kann mit `<em>` betont werden. Die im nächsten Abschnitt besprochenen Links sind selbst ebenfalls Markierungen *innerhalb* von Texten.

Bemerkung: Die meisten Browser versuchen über Fehler bei den Markierungen hinwegzusehen. Damit ein Dokument das größtmögliche Publikum erreichen kann, ist die Einhaltung des Standards ratsam.

## „Hyper“-Text – Links verbinden Dokumente

Links werden mit der Marke `<a>` gekennzeichnet. Hier verwenden wir zum ersten Mal ein Attribut. Das Attribut heißt `href` und sein Inhalt gibt das Ziel des Links an. Betrachten wir uns den Link `<a href="http://www.xyzyzy.de/">Link</a>`. Der Link zeigt auf die Seite `www.xyzyzy.de`. In unserem Dokument sehen wir das Wort *Link*. Möglicherweise ist es unterstrichen oder von anderer Farbe, damit wir sehen, daß wir hier einen Link haben. Auf jeden Fall können wir den Mauszeiger auf das Wort bewegen. Wenn wir jetzt die Maustaste drücken, wird die Seite `www.xyzyzy.de` geladen.

Es gibt verschiedene Arten von Links, absolute und relative. Ein absoluter Link gibt den gesamten Pfad zu einem Dokument an, einschließlich des Rechnernamens, auf dem das Dokument liegt. Absolute Links sind `http://www.xyzyzy.de/`, `http://www.roklein.de/legal.html`.

Relative Links dagegen zeigen auf eine Seite *relativ* zum aktuellen Dokument. Wenn Sie auf der Straße nach dem Weg zum Bahnhof gefragt werden antworten Sie ähnlich: „Gehen Sie *von hier aus* die dritte Straße links.“ Betrachten Sie zum Beispiel gerade die Seite `http://www.roklein.de/`, so ist gibt es dort einen relativen Link: `<a href="legal.html">Impressum</a>`.

## Aufgaben

Schreibe drei HTML-Dokumente. Jedes Dokument soll mindestens zwei Absätze mit Überschriften enthalten. Eines der Dokumente soll Links auf die beiden anderen Dokumente enthalten.

Überlege, ob absolute oder relative Links sinnvoll sind.

Ändere die drei Dokumente dahingehend ab, daß eines zusätzlich eine Aufzählungsliste enthält und eines eine ungeordnete Liste.

Lade das Programm `tidy` von `http://tidy.sourceforge.net/` herunter und installiere es. Man kann es mit `tidy -e <dateiname>` von der Kommandozeile starten, um die Fehler und Probleme in HTML-Dokumenten zu finden. Starte `tidy` mit den drei HTML-Dokumenten. Versuche Fehler und Warnungen zu erklären.

## Tabellen

Information lässt sich oft gut in Tabellenform darstellen. In HTML werden Tabellen oft dazu verwendet, das Aussehen der Seite besser zu bestimmen. Davon sollte man – wenn möglich – absehen.

Eine Tabelle wird in HTML mit `<table>` begonnen und mit `</table>` beendet. Ihr Inhalt besteht aus Zeilen: `<tr>` und `</tr>`; *tr* steht für *table row*. Die Zeilen enthalten Zellen, `<td>` und `</td>` bzw. `<th>` und `</th>`. Die Elementmarker *td* und *th* stehen für *table data* bzw. *table header*. Beide beschreiben eine Datenzelle, wobei *th* Zellen als Zeilen- und Spaltenüberschriften auszeichnet.

Beispiel einer Tabelle:

```
<table>
  <tr>
    <th>&nbsp;</th>
    <th>Januar</th>
    <th>Februar</th>
  </tr>
  <tr>
    <th>Regenmenge</th>
    <td>100ml</td>
    <td>107ml</td>
  </tr>
  <tr>
    <th>Regentage</th>
    <td>17</td>
    <td>13</td>
  </tr>
</table>
```

&nbsp;	Januar	Februar
Regenmenge	100ml	107ml
Regentage	17	13

Es ist möglich, die Tabelle in Kopf-, Fußzeilen und Tabellenkörper einzuteilen. Die Kopf- und Fußzeilen werden beim Drucken auf jeder Seite wiederholt. Die Einteilung erfolgt mit den Marken `<thead>`, `<tfoot>` und `<tbody>` und den zugehörigen Endmarken. Es müssen immer alle drei Marken vorhanden sein. Die Reihenfolge ist festgelegt: zuerst kommt *thead*, dann *tfoot*, dann ein oder mehrere *tbody*.

## Bilder einbinden

Bilder werden mit Hilfe der Marke `<img>` eingebunden. Diese Marke hat *keine* entsprechende Endmarke. Die Quelldatei des Bildes wird als Attribut (siehe auch nächster Abschnitt) von *img* angegeben:

```

```

## Attribute

Viele Marken können mit Attributen versehen werden. Es gibt „Pflicht“-Attribute, wie *src* beim Einbinden von Bildern. Weitere Attribute von *img* sind *width* und *height*, damit werden die Breite und Höhe eines Bildes angegeben. Sind die Attribute vorhanden, kann der Browser das Layout einer Seite festlegen, bevor die Bilder geladen sind; die bereits geladenen Elemente können dargestellt werden, bevor die ganze Seite geladen ist. Ein weiteres Attribut für *img* ist *alt* mit dem ein alternativ darzustellender Text festgelegt wird. Das ist wichtig für Browser, die keine Bilder darstellen können oder den Inhalt der Webseite vorlesen.

Die Werte von Attributen werden immer in doppelten Anführungsstrichen angegeben:

```

```

## Aufgaben

Erstellen Sie ein Dokument mit Bildern und einer Tabelle mit den Marken *thead*, *tfoot*, und *tbody* und dem Tabellenattribut *border*. Informationen dazu und zu anderen Attributen gibt es auch in der *selfhtml*<sup>1</sup>-Dokumentation.

<sup>1</sup> <http://de.selfhtml.org/> – Eine sehr gute und ausführliche Referenz zu HTML

## Markierungen innerhalb von Texten

Innerhalb von Texten gibt es Markierungen wie die schon oben angesprochene Marke `<em>`. Andere sind `<strong>`, die wie *em* den eingeschlossenen Textabschnitt betont. Allgemein wird der von *em* eingeschlossene Text kursiv dargestellt, der von *strong* eingeschlossene fett. Die Marke `<cite>` wird für Zitate oder Verweise zu anderen Quellen verwendet. Wir haben vier Marken, die wir in Verbindung mit Computern verwenden können. Fragmente von Computerprogrammen werden mit `<code>` eingeschlossen. Für Beispiele verwenden wir `<samp>` und mit `<kbd>` markieren wir Tastatureingaben. Weiterhin werden Variablen mit `<var>` markiert. Für Abkürzungen und Akronyme stehen und zwei Marken zur Verfügung, `<abbr>` und `<acronym>`. Bei beiden können wir die Langform mit Hilfe des `title`-Attributs angeben:

```
<abbr title="und so weiter.">usw.</abbr>
<acronym title="The New England Science Fiction Association,
Inc.">NESFA</acronym>
```

## Markierungen auf Blockebene

Außer *p* gibt es einige weitere Markierungen auf Blockebene. Mit `<blockquote>` in einem eigenen Abschnitt dargestellt. Die Marke `<address>` verwenden wir um Adressen – gewöhnlicherweise kurze Adressen, wie Name und Emailadresse – anzugeben. Die Marke `<pre>` leitet vorformatierten Text ein. HTML-Elemente *innerhalb* des vorformatierten Textes werden interpretiert.

## Zeilenumbrüche

Gelegentlich wollen wir einen Zeilenumbruch an einer bestimmten Stelle. Dazu verwenden wir die Marke `<br>`, für die es *keine* Endmarke gibt. Wenn das Dokumentenformat mit XHTML verträglich sein soll, schreiben wir statt `<br>` eine Version mit eingeschlossenem Endemarker: `<br />`.

Wollen wir aber einen Textabschnitt *ohne* Zeilenumbruch, so schließen wir ihn mit `<nobr>` und `</nobr>` ein. Wenn wir aber nur an einem bestimmten Leerzeichen einen Zeilenumbruch verhindern wollen, so verwenden wir `&nbsp;` statt eines Leerzeichens:

Wir wollen keinen Zeilenumbruch: HTML&nbsp;4.01

## Trennlinien

Hin und wieder wollen wir Textabschnitte durch Trennlinien abtrennen. Dafür steht und die Marke `<hr>` zur Verfügung. Wie *br* gibt es keine Endmarke für *hr* und die XHTML-verträgliche Schreibweise ist entsprechend `<hr />`.

## Definitionslisten

Bei den Aufzählungen haben wir bisher eine Form ausgelassen, die Definitionslisten. Im Gegensatz zu den anderen Listen gibt es hier keine *li*-Marke, sondern die Marken `<dt>` und `<dd>`, von denen der zu definierende Begriff (*dt* – definition term) und die Definition selbst (*dd* – definition description) eingeschlossen werden. Definitionslisten sind insbesondere dazu gedacht, Glossare zu erstellen.

## Kommentare

Wir können auch Kommentare im HTML-Code einschließen. Kommentare beginnen mit der Marke `<!--` und enden mit `-->`. Kommentare dürfen mehrere Zeilen lang sein, sie dürfen aber keine Folge von Bindestrichen (Zwei oder mehr Bindestriche hintereinander) enthalten.

## Aufgaben

Schreibe mindestens fünf HTML-Dokumente über ein beliebiges selbstgewähltes Thema, in denen möglichst viele der besprochenen Elemente verwendet werden. Die Dokumente sollen untereinander verlinkt sein.

## Beispiel

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN"
"http://www.w3.org/TR/REC-html40/strict.dtd">
<html>
<head>
  <meta http-equiv="content-type"
  content="text/html; charset=iso-8859-1">
  <title>Die erste Seite</title>
</head>
<body>
  <h1>Meine erste HTML-Seite</h1>
  <p>Das ist meine erste HTML-Seite. Darin finden sich nur
  ein paar Beispieltexte, um einfache HTML-Elemente zu
  erklären.</p>

  <h2>Wieso ein Beispiel</h2>
  <p>Beispiele erlauben es dem Lernenden sich das zu
  Lernende einfacher zu veranschaulichen. Anhand eines
  Beispiels lässt sich etwas leichter nachvollziehen und
  das Beispiel kann für eigene Übungen zur Vorlage
  herangezogen werden.</p>
  <p>Das sich so einstellende Erfolgserlebnis motiviert den
  Lernenden zu einer besseren Leistung.</p>

  <h2>Beispiel für ungeordnete Listen</h2>
  <p>Es gibt verschiedene Sorten von Tee. Es gibt eine sehr
  grobe Unterteilung in vier Sorten:</p>
  <ul>
    <li>grüner (unfermentierter) Tee</li>
    <li>schwarzer (fermentierter) Tee, in China auch
    <em>roter Tee</em> genannt.</li>
    <li>Oolong-Tee, ein leicht anfermentierter Tee</li>
    <li>der selten vorkommende weiße Tee</li>
  </ul>

  <h2>Beispiel für geordnete Listen</h2>
  <p>Kinder lernen eine bestimmte Reihenfolge von
  Tätigkeiten, die sie vornehmen, wenn Sie eine Straße
  überqueren:</p>
  <ol>
    <li>Nach links schauen, ob kein Auto kommt</li>
    <li>nach rechts schauen, ob kein Auto kommt</li>
    <li>nochmal links kontrollieren</li>
    <li>bis zur Mitte gehen</li>
    <li>in der Mitte kurz rechts kontrollieren</li>
    <li>die Straße fertig überqueren</li>
  </ol>

  <table>
    <tr>
      <th>&nbsp;</th>
      <th>1</th>
      <th>2</th>
    </tr>
    <tr>
      <th>1</th>
      <td>1</td>
      <td>2</td>
    </tr>
    <tr>
      <th>2</th>
      <td>2</td>
      <td>4</td>
    </tr>
  </table>
</body>
</html>
```