

(X)HTML - Einführung



ID: TT-VL4
Dokumenten URL: <http://docs.tx7.de/TT-VL4>
Autor: Tom Gries <tom@tx7.de>
Version: 5.0.0 vom 01.07.2017

>> Themen

Was ist (X)HTML

Wesentliche Unterschiede HTML / XHTML

Einführung

Die 5 wichtigsten (X)HTML Regeln

Entities

Aufbau einer (X)HTML Datei

Referenzen

>> Was ist (X)HTML?

- **HTML** ist eine Auszeichnungssprache zur Strukturierung von Inhalten.
- **XHTML** ist eine Seitenbeschreibungssprache zur Strukturierung und semantischen Auszeichnung von Inhalten.
- **HTML5** ist der Nachfolger von HTML und XHTML

>> Wesentliche Unterschiede HTML / XHTML

	HTML	XHTML
Groß-/Kleinschreibung der Elemente und Attributnamen	Nicht relevant = = 	Nur klein erlaubt
Elemente ohne Inhalt (z. B. , , <hr>)	Erlaubt	Nicht erlaubt. Jedes Element muss geschlossen werden. Entweder mit End-Tag (</br>) oder besser mit leerem Tag ()
Elemente ohne End-Tag	Teilweise erlaubt.	Nicht erlaubt. End-Tags müssen immer mit angegeben werden.
Attribute in Anführungszeichen	Nur bei bestimmten Zeichen nötig	Immer
Attribute ohne Wertzuweisung	Möglich, bzw. erlaubt <input type="radio" checked>	Nicht erlaubt, z. B. <input type="radio" checked="checked" />

Tags (Auszeichnungen)

- Ein Tag ist eine Markierung.
- Den Begriff Tag kennt man nicht nur in (X)HTML.
- Ein anderes Anwendungsbeispiel für Tags sind MP3 Dateien. Diese werden auch mit Tags versehen (z. B. Autor, Titel, Jahr usw.).

>> Einführung

- In (X)HTML gibt es zwei Arten von Tags:
 - öffnende (Start)
 - schließende (End)
- Tags beginnen mit einem "kleiner als" (<) Zeichen. Dann folgt der Name des Tags. Beendet werden Tags mit einem "größer als" (>) Zeichen.
- Beim schließenden Tag kommt vor dem Namen noch ein "Slash" (/).
- Tags bestehen aus einem Wort ohne Sonderzeichen.
Beispiel: Start Tag:= <tag> End Tag:= </tag>

Attribute

- Attribute erweitern Tags um Eigenschaften, z. B. Größenangaben oder Formatierungsangaben.
- Attribute bestehen aus einem Namen und einer Wertzuweisung.
- Die Zuweisung erfolgt durch das '=' Zeichen.
Beispiel: `type="radio"`

>> Einführung

- In einem Tag sind mehrere Attribute zulässig. Diese werden dann durch Whitespaces getrennt notiert.
- Attribute stehen immer im öffnenden Tag und werden nach dem Tagnamen notiert.
- Die Reihenfolge kann frei gewählt werden.

Beispiel:

Start Tag: `<tag attribut1="10" attribute2="blue">`

End Tag: `</tag>`

Elemente

- Als Element wird alles vom öffnenden bis zum schließenden Tag inklusive dem Text zwischen den Tags bezeichnet.

Beispiel für ein Element:

```
<tag attribut="foo">Ein bisschen Text</tag>
```

- Elemente werden in Block-Elemente und Inline-Elemente unterteilt.

>> Einführung

- Block-Elemente erzeugen einen Abstand (z. B. einen Zeilenabstand) zwischen dem vorherigem und dem nächsten Element (z. B. <p>). Inline-Elemente erzeugen keinen Abstand (z. B.).
- Block-Elemente können andere Block-Elemente oder Inline-Elemente enthalten.
- Inline-Elemente enthalten nur andere Inline-Elemente.

Entities

- Entities sind Akronyme für Zeichen, die nicht direkt in HTML eingegeben werden können oder dürfen. Man kann ein Entity daher als Alias für das Zeichen auffassen.
- Entities beginnen mit einem & und werden mit einem ; abgeschlossen.

Beispielentity für das Euro-Zeichen: € (= €)

Numeric Character Referenz

- Unter Numeric Character Referenz versteht man die Darstellung von Zeichen in numerischer Schreibweise.
- Man unterscheidet dezimale und hexadezimale Numeric Character Referenz.
- Im Gegensatz zu den Entities lassen sich mit der numerischen Darstellung alle Zeichen abbilden. Ein HTML Dokument könnte also komplett mit numerischen Referenzen geschrieben werden.

>> Einführung

- Dezimale numerische Referenzen beginnen mit &#
 - Hexadezimale numerische Referenzen mit &#x
- Beide werden mit einem Semikolon (;) abgeschlossen.

Beispiel:

DEC: **
**

HEX: **
**

Zeichensatz / Zeichensatztabellen

- Weil das Internet international ist und jede Sprache eigene Zeichen hat sind auch unterschiedliche Zeichensatztabellen erforderlich.
- Der einzige Zeichensatz, der in jeder Sprache und für jedes Betriebssystem existiert, ist der ASCII Code.
- Der ASCII Code besteht aus 128 Zeichen. Der ANSI- und DOS-Zeichensatz ist eine Erweiterung des ASCII Codes auf 256 Zeichen.

➤ Einführung

HEX	x0	x1	x2	x3	x4	x5	x6	x7	x8	x9	xA	xB	xC	xD	xE	xF
0x	NUL	SOH	STX	ETX	EOH	ENQ	ACK	BEL	BS	HAT	LF	VT	FF	CR	SO	SI
1x	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2x	SP	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3x	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4x	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5x	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6x	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7x	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

DEC	x0	x1	x2	x3	x4	x5	x6	x7	x8	x9		x0	x1	x2	x3	x4	x5	x6	x7	x8	x9
0x	NUL	SOH	STX	ETX	EOH	ENQ	ACK	BEL	BS	HAT	1x	LF	VT	FF	CR	SO	SI	DLE	DC1	DC2	DC3
2x	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	3x	RS	US	SP	!	"	#	\$	%	&	'
4x	()	*	+	,	-	.	/	0	1	5x	2	3	4	5	6	7	8	9	:	;
6x	<	=	>	?	@	A	B	C	D	E	7x	F	G	H	I	J	K	L	M	N	O
8x	P	Q	R	S	T	U	V	W	X	Y	9x	Z	[\]	^	_	`	a	b	c
10x	d	e	f	g	h	i	j	k	l	m	11x	n	o	p	q	r	s	t	u	v	w
12x	x	y	z	{		}	~	DEL													

>> Die 5 wichtigsten (X)HTML Regeln

Regel 1:

Ein Tag steht immer in spitzen Klammern.

→ Spitze Klammern können im Text nicht literal dargestellt werden.

Abhilfe:

Darstellung als Entity: < und >

>> Die 5 wichtigsten (X)HTML Regeln

Regel 2:

Ein Tag kann Attribute enthalten. Attribute bestehen aus dem Namen des Attributes und einer Wertzuweisung (in Anführungszeichen). Die Anführungszeichen kennzeichnen den Bereich von Wertzuweisungen.

→ Anführungszeichen dürfen im Text nicht literal dargestellt werden.

Abhilfe:

Darstellung als Entity: "

>> Die 5 wichtigsten (X)HTML Regeln

Regel 3:

Entities werden mit einem '&' eingeleitet und mit einem ';' beendet.

→ Das '&' darf im Text nicht literal verwendet werden.

Abhilfe:

Darstellung von '&' als Entity: &

>> Die 5 wichtigsten (X)HTML Regeln

Regel 4:

Für (X)HTML Dokumente darf im Source-Code ohne Angabe des zu verwendenden Zeichensatzes (Character Encoding) nur der ASCII Zeichensatz verwendet werden.

Abhilfe:

HTML 4.01

```
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
```

HTML 5

```
<meta charset="UTF-8">
```

>> Die 5 wichtigsten (X)HTML Regeln

Regel 5:

Mehrere Whitespaces (Leerzeichen, Tabulatoren oder Zeilenumbrüche) hintereinander werden vom Browser zu **einem** Leerzeichen zusammengefasst.

Abhilfe:

Sind Leerzeichen und/oder Zeilenumbrüche in der Browseransicht gewünscht, müssen diese mit Entities oder Tags (` ` / `
`) im Source-Code notiert werden.

>> Entities

Basic Entities (aus ASCII Zeichensatz):

Darzustellendes Zeichen	Entity
<	<
>	>
&	&
"	"

Basic Entities sind sprachunabhängig und müssen **immer** verwendet werden (siehe Regeln).

Die 16 gebräuchlichsten Entities:

Latin-1 (10)	
 	
ä	ä
Ä	Ä
ö	ö
Ö	Ö
ü	ü
Ü	Ü
ß	ß
©	©
®	®

Basic (4)	
<	<
>	>
&	&
"	"

Symbol (1)	
™	™

Special (1)	
€	€

Weitere Entities:

- Latin-1:
<http://docs.tx7.de/TT-RJ2>
- Symbolische und griechische Zeichen:
<http://docs.tx7.de/TT-4U4>
- Sonderzeichen:
<http://docs.tx7.de/TT-JCQ>

➤➤ Aufbau einer (X)HTML Datei

Das Grundgerüst	
<html>	1 Mal
<head>	1 Mal
<title>	1 Mal
<meta>	0 - n Mal
<body>	1 Mal
restliche Tags	0 - n Mal

>> Aufbau einer (X)HTML Datei

Um Probleme in (X)HTML und zukünftigen Versionen zu vermeiden gelten folgende Regeln:

– Tags werden immer klein notiert:

<table>** anstatt **<TABLE>** oder **<Table>

– Wertzuweisungen bei Attributen stehen immer in Anführungszeichen:

<td width="20">** anstatt **<td width=20>

➤➤ Aufbau einer (X)HTML Datei

- Tags werden immer geschlossen:

<p>Text</p>

anstatt <p>Text

Item

anstatt Item

*
*

*anstatt
*

<hr />

anstatt <hr>

** *anstatt *

>> Aufbau einer (X)HTML Datei

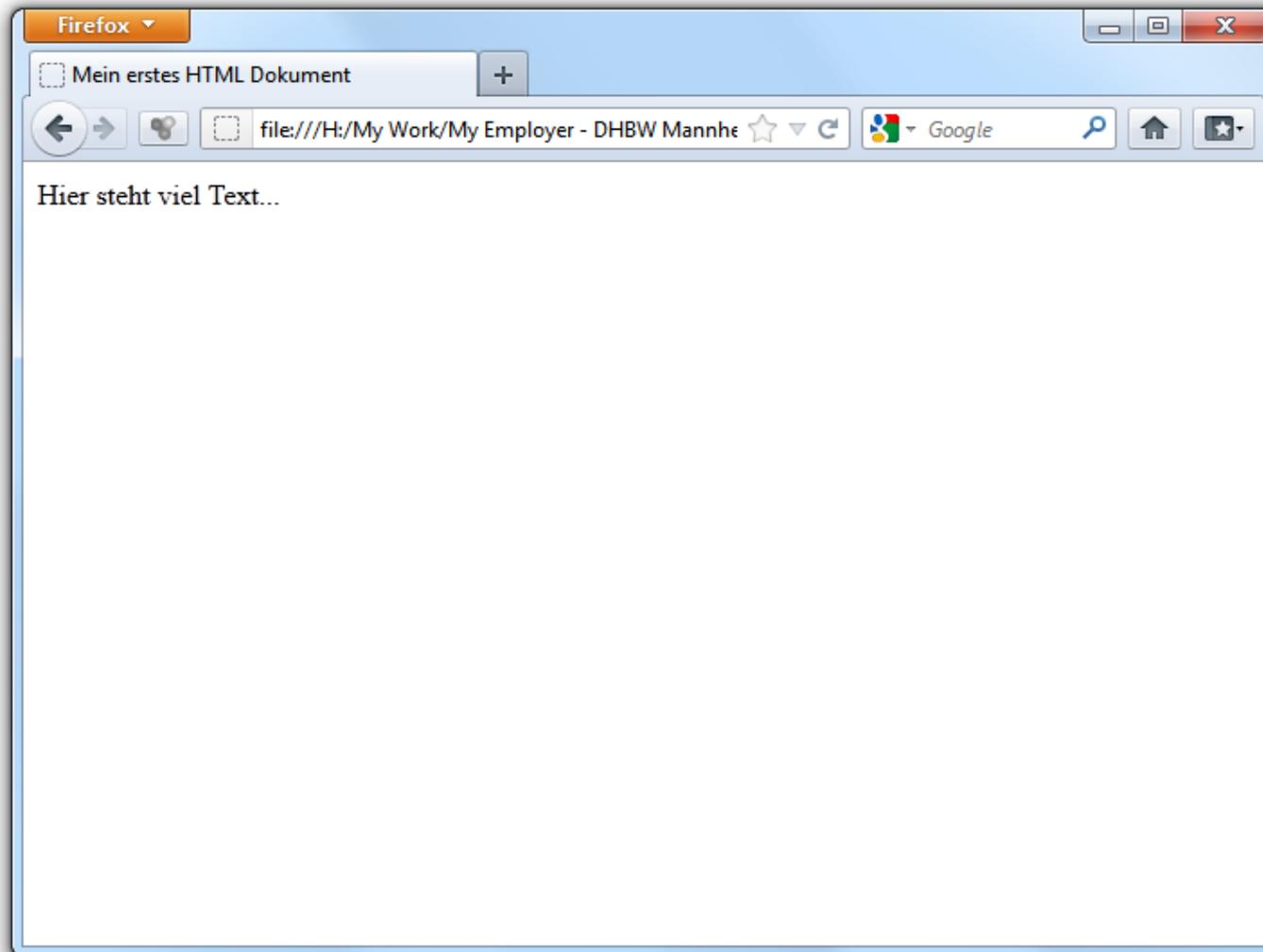
Minimales HTML Dokument ohne DocType (Beispiel 1):

```
<html>  
  <head>  
    <title>Mein erstes HTML Dokument</title>  
  </head>  
  <body> Hier steht viel
```

Text...

```
  </body>  
</html>
```

➤➤ Aufbau einer (X)HTML Datei



>> Aufbau einer (X)HTML Datei

Kommentare:

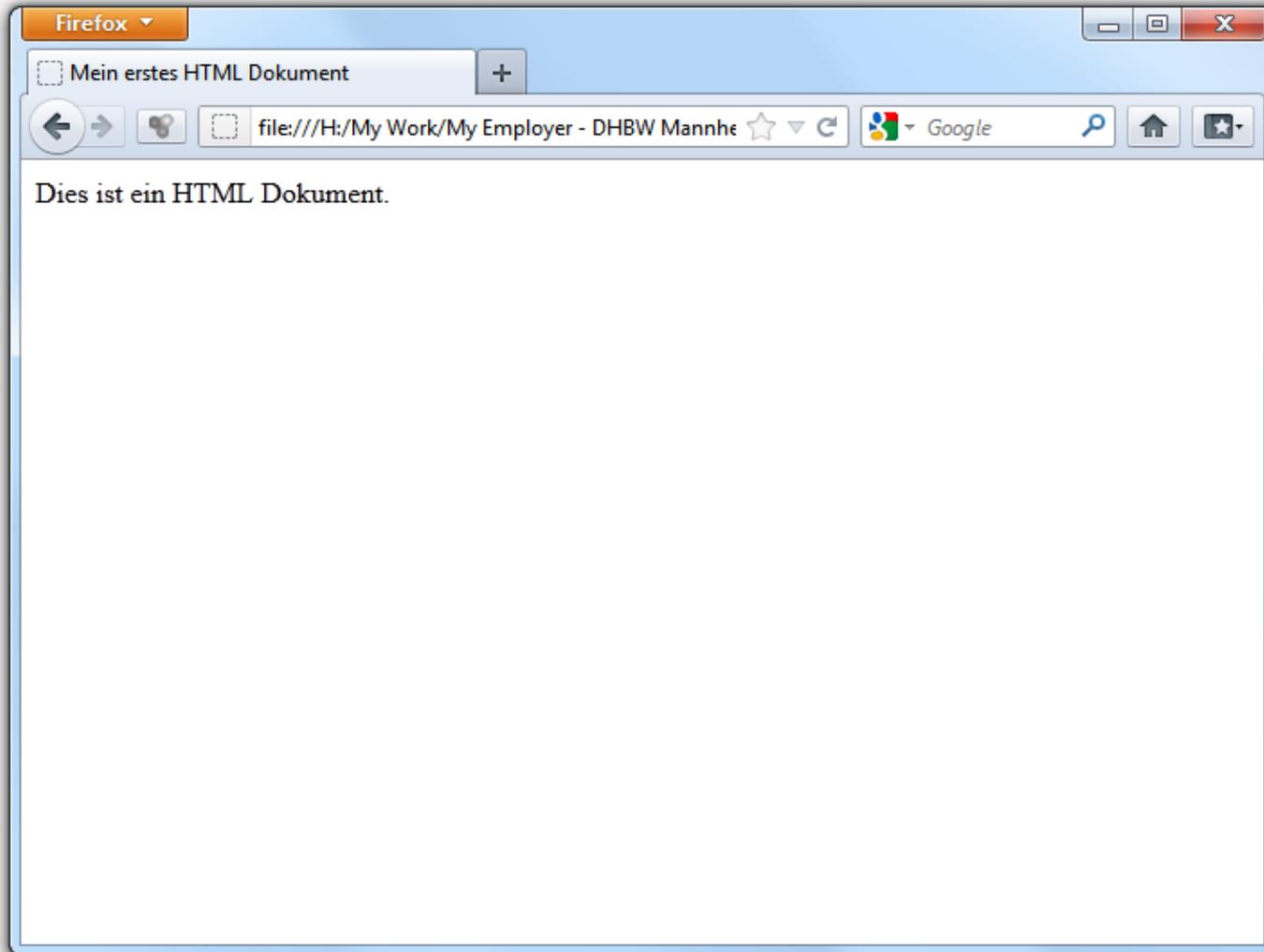
- Kommentare werden mit `<!--` eingeleitet und mit `-->` beendet.
- Alles zwischen den Kommentarmarkierungen (Tags und Fließtext) wird vom Browser ignoriert.

>> Aufbau einer (X)HTML Datei

Kommentare können an (fast) beliebiger Stelle eingefügt werden (Beispiel 2):

```
<html>
  <head>
    <title>Mein erstes HTML Dokument</title>
  </head>
  <body>
    Dies ist ein <!-- einfaches --> HTML
    Dokument.
  </body>
</html>
```

➤➤ Aufbau einer (X)HTML Datei



>> Aufbau einer (X)HTML Datei

Die wichtigsten Tags:

<html>

<head>

<title>

<body> HTML Grundgerüst

<meta> HTML Kopfdaten

<h[1-6]> Heading

<p> Paragraph

 Line break

>> Aufbau einer (X)HTML Datei

Die wichtigsten Tags:

- `` Aufzählungsliste
Sub-Tag(s): ``
- `` Numerierte Liste
Sub-Tag(s): ``
- `<dl>` Definitionsliste
Sub-Tag(s): `<dt>`, `<dd>`
- `<hr>` Trennlinie

>> Aufbau einer (X)HTML Datei

Die wichtigsten Tags:

, <i>, Physische Text-
<u> auszeichnungen

<a> Verweise (Links

 Grafiken

<table> Tabellen

Sub-Tag(s): <th>¹, <tr>¹, <td>¹

¹ Nur eine Auswahl - weitere Sub-Tags in den SelfHTML Erläuterungen

>> Aufbau einer (X)HTML Datei

Die wichtigsten Tags:

<form> Formulare

Sub-Tag(s): <input>², <textarea>², <select>²

<frameset> Frames

Sub-Tag(s): <frame>, <noframes>

² Nur eine Auswahl - weitere Sub-Tags in den Erläuterungen

>> Referenzen

SelfHTML:

<http://docs.tx7.de/TT-EQ3>

HTML Kurse:

<http://docs.tx7.de/TT-94S>