

HTML - HyperText Markup Language

Gute Online-Unterlagen zum Selbstlernen:

https://wiki.selfhtml.org/ resp. https://selfhtml.org/

https://www.w3schools.com/

https://developer.mozilla.org/de/docs/Learn

HyperText Markup Language



HTML ist die Beschreibungssprache des WWW

- Hypertext: Dokumente, die Verweise auf andere Dokumente enthalten
- universelle, plattformunabhängige Markup-Sprache
- Markup steht dabei für Auszeichnung und bedeutet, dass Teile eines Textes als besondere strukturelle Elemente ausgezeichnet werden können. Beispiele: Überschriften, Tabellenstrukturen, ...
- Bei HTML wird die Auszeichnung mit einer Semantik verknüpft
- Auszeichnungselemente können einen Inhalt haben

Eigenschaften von HTML

19

- HTML dient der Definition des Aufbaus einer Internetseite: Elemente, Strukturen, Verweise (Hyperlinks), referenzierte Elemente (Grafiken, Multimedia...)
- Seit 1992 standardisiert und zunehmend fortentwickelt
- Im Jahr 2014 wurde HTML 5 standardisiert

HyperText Markup Language



HTML, und was gibt es noch?

HTML: Inhalt und Aufbau/Grundstruktur der Seite

CSS: Design/Erscheinungsbild des Dokumentes

JavaScript: Logik, Programmierung beim Verarbeiten des Dokumentes





Hierarchische Gliederung der Auszeichnungselemente, z.B.

- Auszeichnungselemente haben einen festen Erstreckungsraum
- Sie haben eine Semantik und ggf. einen Inhalt
- Elemente können verschachtelt sein
- Verschachtelte Elemente müssen vollständig in den Erstreckungsraum der übergeordneten Elemente enthalten sein
- Repräsentation als Baum möglich (DOM-Baum)

```
[Überschrift] Text der Überschrift [Ende Überschrift]
[Liste]
[Listenpunkt] Text des Listenpunkts [Ende Listenpunkt]
[Listenpunkt] Text des Listenpunkts [Ende Listenpunkt]
[Ende Liste]
```

HTMLAuszeichnungselemente: Tags



HTML-Elemente werden durch sogenannte "Tags" markiert

- Tags werden in spitzen Klammern (<, >) notiert
- Die meisten HTML-Elemente werden durch ein einleitendes (<>) und ein abschließendes (</>) Tag markiert, außer der Erstreckungsraum ist auch ohne schließendem Tag klar
- Mit diesen Elementen ist über den Erstreckungsraum ein Inhalt assoziiert, der auch weitere Auszeichnungselemente beinhalten kann. Hier spielt die Reihenfolge eine Rolle!
- Es gibt Elemente, sogenannte empty-Elements, die grundsätzlich keinen Inhalt haben und von daher keinen schließenden Tag benötigen (Beispiel:
, <col>, <link>,)

https://developer.mozilla.org/de/docs/Glossary/empty_element

 Tags können mittels Attributen (Key-Value-Paaren) präzisiert werden (Notation: name="wert"). Diese sind spezifisch zum Tag. Die Reihenfolge spielt hierbei keine Rolle

HTMLAuszeichnungselemente: Tags

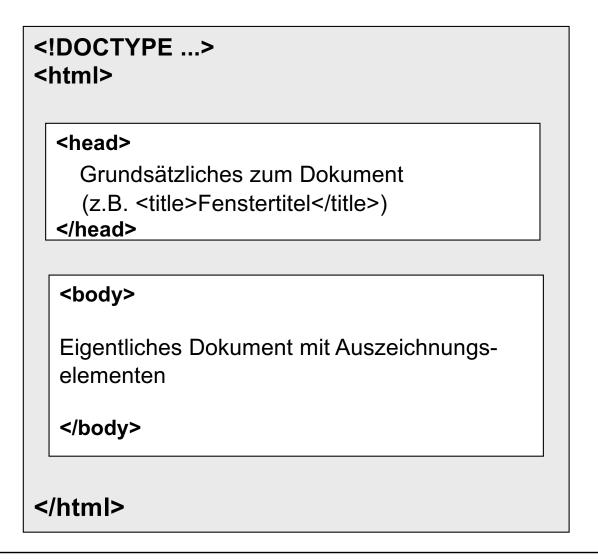


```
<h1>Überschrift 1. Ordnung</h1>
<strong>Wichtiger Text in Fettschrift</strong>
Ein Zeilenumbruch wird durch den Tag <br/>
eingeleitet
<h1>HTML - die Sprache des Web</h1>
<img src="katze.jpg" alt="Katze" width="100" height="100">
```

HTMLGrundlegender Aufbau



Vereinfachte Darstellung eines HTML-Dokuments



HTML Grundlegender Aufbau



```
<!DOCTYPE html>
                                 Kommentar, keine HTML-Auszeichnung
<html>
                                  Wird als (empfohlene!) Anweisung
                                  für den Browser genutzt, der weiß, was
<head>
                                  zu verarbeiten ist
    <meta charset="utf-8">
    <title>Titel der Webseite</title>
</head>
<body>
                                       Der Inhalt des h1-
<h1>Überschrift</h1>
                                       Elements ist "Überschrift"
Text mit Zeilenumbruch am Ende<br>
An dieser Stelle erfolgt kein Umbruch:
Sag ich doch!
</body>
</html>
```

HTML

Darstellung als DOM-Baum



Darstellung des HTML-Dokumentes als DOM-Baum

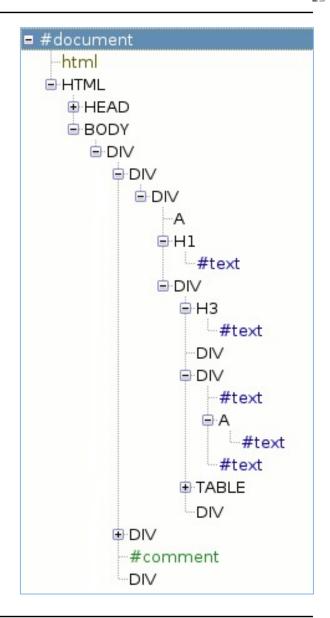
- Darstellungsform bei der Verarbeitung
- Tags und Elementinhalte werden durch Knoten repräsentiert
- Werkzeuge zum Navigieren durch den Baum erlauben Zugriff auf alle Tags und Attribute der Webseite

Beispiel (rechte Seite)

- Teilweise ausgeklappter DOM-Baum
- Text ist ein eigener DOM-Knoten (#text), der den Inhalt repräsentiert
- Reihenfolge wichtig

Tag-Attribute

 Zusätzlich kann jeder Knoten Eigenschaften durch Attribute erhalten



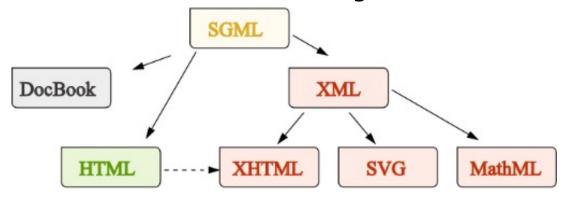


XML (Extensible Markup Language) und HTML

- HTML ursprünglich unabhängig von XML entwickelt
- Beide entstanden aus Standard Generalized Markup Language (SGML)
 - > Ähnlicher Aufbau (Tags, Baumstruktur)
- XML erlaubt beliebige Tags, HTML hat vorgegebene Tags
- HTML-Auszeichnungen implizieren eine Semantik, z.B. Überschrift

XHTML

- Vereinigung von HTML und XML
- Version 2.0 zu Gunsten von HTML 5 eingestellt



Quelle: https://wiki.selfhtml.org/wiki/Datei:Auszeichnungssprachen.svg

ISO 8879 (SGML): https://www.iso.org/standard/16387.html



HTML 4 (1997), HTML 4.01 (1999)

- Teils mit unspezifischen Vorgaben
- Einfach HTML zu schreiben
- Schwierig browserübergreifend das gleiche Layout zu erzeugen

XHTML: Extensible HTML (~2000)

- Strikter Standard sollte die Fehler von HTML 4 beheben
- Durchsetzung von XHTML scheitert auf Grund der Striktheit
 - > Minimale Fehler im XML können zum kompletten Ausfall der Seite führen, der Browser zeigt einen XML-Parse-Error

HTML 5 (Entwicklung seit ca. 2008, Recommendation 2014)

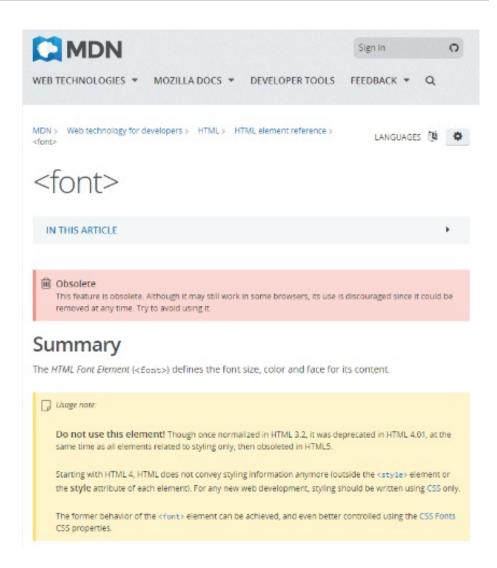
- Ein browserübergreifender einheitlicher Parser soll die Unterschiede im Anzeigen verhindern und gleichzeitig Flexibilität erlauben
- Zahlreiche moderne HTML-Erweiterungen (video, audio)
- Viele "Altlasten" entfernt
- Oft als Sammelbegriff für moderne Technologien genutzt
 Unsere Beispiele sind HTML 5

Weitere Informationen: https://html.spec.whatwg.org/#history-2



HTML Altlasten

- Styling im HTML-Code
 - Veränderung von Farben, Textgrößen, Abständen, usw. im HTML (z.B. mittels font-Tags)
 - > Sollte seit HTML 4 (1997) mittels **CSS** gemacht werden!
- Escaping von Umlauten (und anderen Zeichen)
 - > Bei der Wahl des richtigen Zeichenstatzes ist das nicht nötig
- JavaScript im HTML-Code
 - > HTML, JavaScript und CSS sollten durch eigene Dateien getrennt sein



Quelle: https://developer.mozilla.org/en-US/docs/Web/HTML/Element/font



Beispiele für schlechten Code

Browserunterstützung

 Trotz Auszeichung als "obsolete" unterstützen moderne Browser aus Gründen der Kompatibilität i.d.R. immer noch alle veralteten Tags und Attribute

Viele Internetseiten enthalten immer noch schlechte Codebeispiele

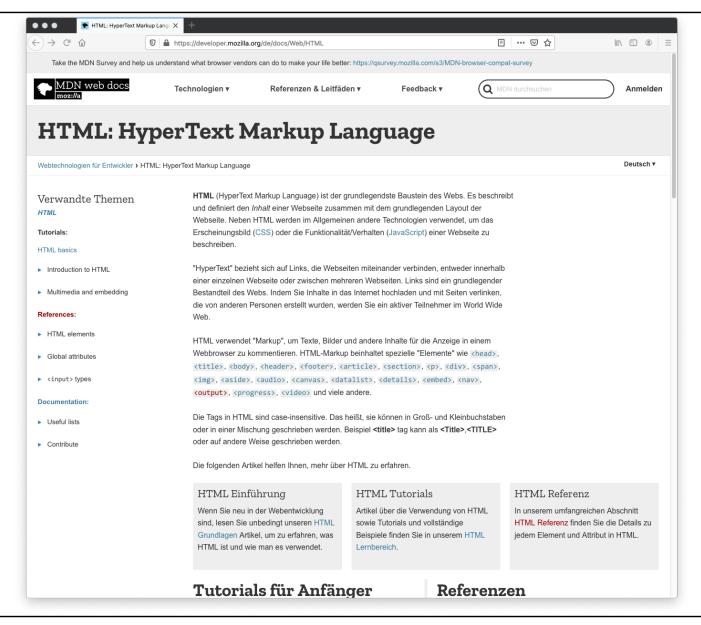
Im Zweifel MSDN oder MDN aufsuchen

Weitere Informationen: https://www.w3.org/TR/html5/obsolete.html

HTMLMozilla Developer Network







HTML Zeichensatz



Zeichensatz des Dokumentes

Bestimmt welche Zeichen für die übermittelten Bytes angezeigt werden

Beispiele:

> ASCII-Byte: 65 (alternativ: 0x41 oder 01000001)

Entspricht: A

> UTF-8 Byte: 65 (alternativ: 0x41 oder 01000001)

Entspricht: A

- Die ersten 128 Zeichen (0-127) des ASCII und UTF-8-Standards sind gleich!
 - > Tatsächlich besteht ein ASCII-Code nur aus 7 Bits

UTF-8 ist abwärtskompatibel zu ASCII

HTML Zeichensatz



UTF-8 ist abwärtskompatibel zu ASCII

 Problem: Es fällt evtl. nicht auf, wenn ein Dokument mit einem falschen Zeichensatz ausgeliefert wurde.

Mögliche Lösung: Andere Zeichen kodieren

- Beispiele:
 - > € statt €
 - > ä statt ä
 - © statt ©
- Unsaubere Lösung!
 - > Vor allem historisch bedingt, da Browser sich in manchen Fällen über den übermittelten Zeichensatz hinwegsetzten
 - > Besser: Korrekten Zeichensatz angeben!
 - > Im <head>-Element: <meta charset="UTF-8">

When not to use escapes: https://www.w3.org/International/questions/qa-escapes#not

HTML Zeichensatz UTF-8



Nur wenige Zeichen müssen tatsächlich "escaped" werden:

- Verwendung eckiger Klammern ohne Interpretation als Auszeichnung > Beispiel: Ein h1-Tag: <h1>
- "Ampersand", das sonst als Einleitung eines Escape-Characters interpretiert wird
- Anführungszeichen innerhalb von Attributen
 - > Beispiel: title="Er sagte " Hallo " ."

Zeichen	Bedeutung	HTML-Code
<	Öffnende Klammer (Auswertung als Tag verhindern)	<
>	Schließende Klammer	>
&	"Ampersand", Käufmännisches Und	&
11	Doppelte Anführungszeichen	"
ı	Einfaches Anführungszeichen	' ;

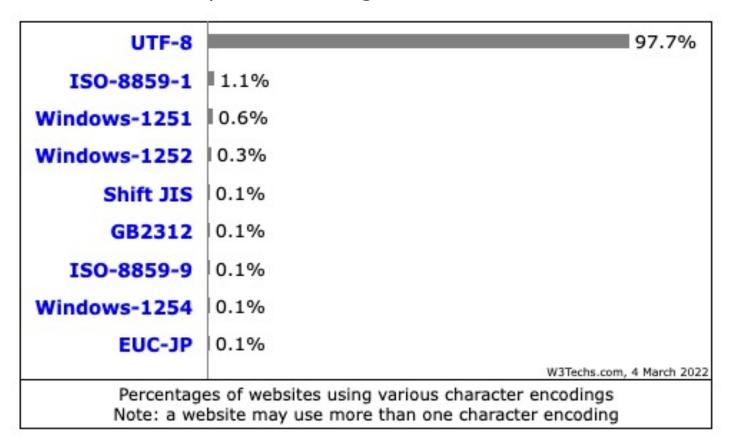
When to use escapes: https://www.w3.org/International/questions/qa-escapes#use

HTML Zeichensatz



Den korrekten Zeichensatz wählen

- Immer UTF-8 wählen!
 - > Weitverbreitetste Implementierung von UNICODE-Zeichen



Quelle: https://w3techs.com/technologies/overview/character_encoding/

HTML Zeichensatz



UTF-8 nutzen

- 1. Editor ggf. auf UTF-8 stellen
- 2. Encoding im HTML angeben

<!DOCTYPE html>

- 3. Prüfen, welche Einstellungen der Server übermittelt
 - > Einstellung im Apache:
 - httpd.conf / .htaccess: AddDefaultCharset UTF-8 oder AddCharset UTF-8
 - > Bei widersprüchlichen Angaben nutzt der Browser das Server-Encoding HTTP-Response Header: Content-Type: text/html;charset=utf-8

Weitere Informationen: https://www.w3.org/International/questions/qa-html-encoding-declarations





Enthält Informationen über das Dokument

Beispiel

- Die Informationen im <head> sind größtenteils für eine automatisierte Verarbeitung von Bedeutung (Browser, Bots)
- Oftmals werden hier auch wesentlich Ressourcen wie CSS und JavaScript eingebunden





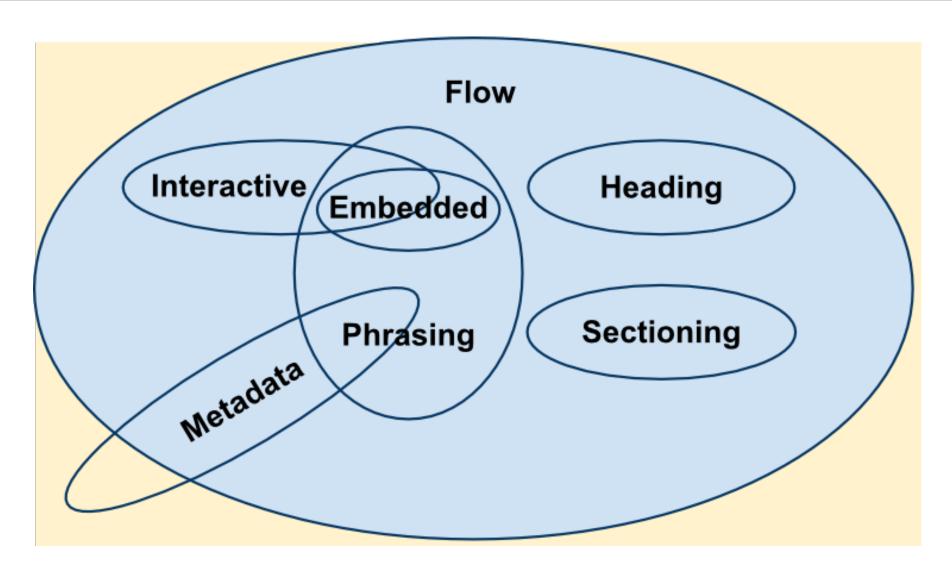
Enthält den eigentlichen Teil des Dokuments

Beispiel

```
<body>
<h1>Überschrift</h1>
<h2>Tabelle</h2>
Land
    BIP
  Deutschland
    3,73 Billionen USD
  UK
    2,81 Billionen USD
  </body>
```

HTML Elemente





Vollständige Liste: https://developer.mozilla.org/en-US/docs/Web/Guide/HTML/Content categories

HTML Elemente Eine Auswahl



Metadaten

<meta>, <title>, <base>, <link>, <script>, <style>

Flow/Fluss-Elemente

<a>, <article>,
, <button>, <form>, <div>, <header>, , <input>, <nav>, , , <script>, <h[1-6]>, <section>,

Elemente zur Aufteilung (HTML5 Semantics-Tags)

<article>, <aside>, <nav>, <section>

Überschriften

<h1>, <h2>, ...

Stil-Elemente

, <button>, <code>, <input>, <label>, , <time>

Elemente zur Einbettung von Inhalten

<audio>, <canvas>, <iframe>, , <object>, <svg>, <video>

Interaktive Elemente

<a>, <button>, <iframe>, <label>, <select>, <textarea>

Vollständige Liste: https://developer.mozilla.org/en-US/docs/Web/Guide/HTML/Content_categories

HTML Elemente Logisch vs. Physisch



Wichtige Unterscheidung

- In HTML4 gab es logische und physische Elemente zur Auszeichnung des Inhaltes (Text)
- Physische Elemente definierten direkt das Aussehen eines Elementes (z.B. unterstrichen, kursiv), treffen aber keine Aussage über die Bedeutung
- Logische Elemente machten eine Aussage über die Bedeutung des Elementes (z.B. um was für Text es sich handelt), die Formatierung hängt vom Browser ab, kann aber auch mit Hilfe von CSS verändert werden

Physische Auszeichnungselemente wurden faktisch mit HTML5 verworfen (auch wenn die Tags noch unterstützt werden)

- Faktisch sollte man möglichst mit logischen Auszeichnungen arbeiten
- Mit CSS kann das Aussehen beliebig verändert werden

HTML Elemente Logisch vs. Physisch



Beispiele für logische Auszeichnungen

```
<strong>stark betont</strong>
<em>betont (emphasized) </em>
<code>Quellcode</code>
<samp>Beispiel</samp>
<cite>Zitat</cite>
<dfn>Definition</dfn>
<abbr>abgekürzte Schreibweise</abbr>
```

Beispiele für alte physische Auszeichnungen

```
<b>fett</b>
<i>kursiv (italic)</i>
<u>unterstrichen (underlined)</u>
<strike>durchgestrichen</strike>
<big>größer als normal</big>
<small>kleiner als normal</small>
<sup>hochgestellt</sup>
<sub>tiefgestellt</sub>
```

Weitere Informationen: https://developer.mozilla.org/de/docs/Web/HTML/Element/strong#Bold_vs. Strong

HTML Elemente Überschriften



HTML unterscheidet 6 Überschriftenebenen

- <h1>Überschrift 1. Ordnung</h1>
- <h2>Überschrift 2. Ordnung</h2>
- <h3>Überschrift 3. Ordnung</h3>
- <h4>Überschrift 4. Ordnung</h4>
- <h5>Überschrift 5. Ordnung</h5>
- <h6>Überschrift 6. Ordnung</h6>

Nicht zu verwechseln mit: <header>...</header>

- Keine Überschrift!
- Element mit Einführungsinformationen zur Navigationshilfe
 - > Logo der Seite
 - > Gruppierung von Überschriften
 - > Autor
- <footer>...</footer> entspricht dem Gegenstück dazu

Weitere Informationen: https://developer.mozilla.org/docs/Web/HTML/Element/h1-h6, https://developer.mozilla.org/docs/Web/HTML/Element/h1-h6, https://developer.mozilla.org/docs/Web/HTML/Element/h1-h6, https://developer.mozilla.org/docs/Web/HTML/Element/header

HTML Elemente Textabsätze und Zeilenumbrüche



Textabsatz

Repräsentiert einen Absatz

```
Erster Absatz.
Hier beginnt der zweite Absatz.
Und hier ist noch ein Absatz.
```


- Erzwingt einen Zeilenumbruch
- Dies sollte nicht als Ersatz für gesehen werden

Weitere Informationen: https://developer.mozilla.org/docs/Web/HTML/Element/p, <a href="https://developer.mozilla.org/docs/Web/HTML/El

HTML Elemente Listen



Unsortierte Listen (Unordered List)

```
      Apfel
      Birne
      Kiwi
```

- Apfel
- Birne
- Kiwi

Sortierte Listen (Ordered List)

```
      Apfel
      Birne
      Kiwi
```

- 1. Apfel
- 2. Birne
- 3. Kiwi

Weitere Informationen: https://developer.mozilla.org/docs/Web/HTML/Element/ol, <a href="https://develo

HTML Elemente Tabellen



Formatierung

- um die Tabelle einzuleiten
- beginnt eine Zeile (table row)
- beginnt eine Header-Zelle
- > beginnt eine normale Zelle (table data)

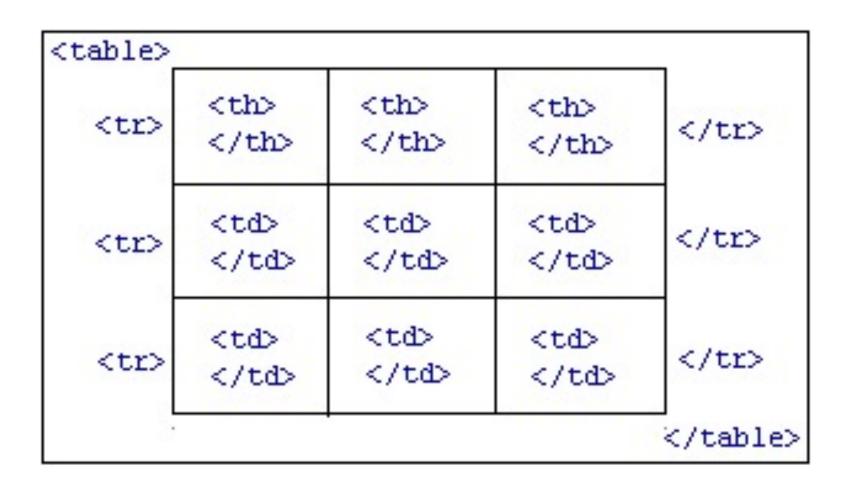
Beispiel

```
\langle t.r \rangle
      Land
      BIP (in Billionen USD) 
  Land
                                          BIP (in Billionen USD)
  Deutschland
                                   Deutschland 3.73
      3,73
                                   UK
                                          2.81
  \langle t.r \rangle
      UK
      2,81
```

Weitere Informationen: https://developer.mozilla.org/en-US/docs/Web/HTML/Element/table

HTML Elemente Tabellen





HTML Elemente Tabellen



Formatierung

- Es ist zusätzlich möglich, dass eine Zelle mehrere Spalten oder Zeilen umfasst
- colspan="2" zum horizontalen Umfassen von zwei Zellen
- rowspan="2" zum vertikalen Umfassen von zwei Zellen

Beispiel (mit CSS-Klassen)

```
        1

        </t
```



HTML Elemente Bilder



Bilder einbinden

-
 - > Das Attribut "src" enthält den (relativen) Pfad (die URL) zur Bilddatei
 - > Das Attribut "alt" enthält einen Alternativtext, der angezeigt wird, wenn das Bild nicht geladen werden kann oder von Screenreadern (barrierefrei) verarbeitet wird
- Die meist unterstützten Grafikformate sind
 - > JPEG, GIF, PNG, BMP, SVG
 - https://en.wikipedia.org/wiki/Comparison_of_web_browsers#Image_form at_support
 - > Empfehlungen:
 - JPEG für Fotos, Bilder mit vielen Farben oder Farbverläufen
 - GIF für Animationen
 - PNG für Logos, kleinere Bilder oder Bilder mit wenig Farben

Beispiel

Weitere Informationen: https://developer.mozilla.org/docs/Web/HTML/Element/img

HTML Elemente Hyperlink



Verbindung zwischen zwei Dokumenten per Link

- Relative und absolute Angaben sind erlaubt
- Beispiele

```
<a href="ziel.html">Verweistext</a>
<a href="https://example.com/datei.html">Ein absoluter Link</a>
<a href="mailto:test@example.com">Email versenden</a>
```

Verweise innerhalb des Dokumentes

- Ein Anker innerhalb des Dokumentes kann auch als Ziel eines Links angegeben werden
- Das Browserfenster springt dann an die jeweilige Stelle
- Beispiel

```
<h2 id="Anker">Überschrift</h2>
<a href="#Anker">Link zur Überschrift</a>
```

Weiterführend: https://developer.mozilla.org/docs/Web/HTML/Element/a, http://w3c.github.io/html/browsers.html#navigating-to-a-fragment-identifier

HTML ElementeGenerische Elemente



<div>...</div>

- Generischer Container, der nichts repräsentiert
- Häufig genutzt um Elemente zu gruppieren oder zu stylen (per CSS)
- Wenn semantisch aussagekräftigere Alternativen verfügbar sind, sollten diese verwendet werden:

```
<nav>, <header>, <footer>, <article>, ...
```

 "block"-Element, d.h. der Container beginnt in einer neuen Zeile, was i.d.R. zu einem Zeilenumbruch führt

...

- Generischer Container wie div
- "inline"-Element, d.h. der Container kann auch innerhalb von Text genutzt werden (kein Zeilenumbruch)
- Beispiel (mit CSS-Klasse highlight)

```
<span class="highlight">Ich</span> bin hervorgehoben.
```

Weitere Informationen: https://developer.mozilla.org/docs/Web/HTML/Element/div, https://developer.mozilla.org/docs/Web/HTML/Element/span

HTML Elemente Formulare



Übermittelung von Daten

 Ein Formular enthält Eingabefelder, die das Versenden von Informationen an einen Server erlauben

Beispiel

<form>...</form>

- action="..." Wohin sollen die Daten gesendet werden
- method="..." Methode zur Übermittlung der Daten
 - > POST oder GET (wird noch näher behandelt)

Weitere Informationen: https://developer.mozilla.org/docs/Web/HTML/Element/form

HTML Elemente Formulare



<textarea>...</textarea>

- Erlaubt die Eingabe von längerem und mehrzeiligem Text
- cols-Attribut bestimmt die Breite, rows die Höhe
 - > Besser über CSS festlegen
- Beispiel

```
<textarea name="textarea" rows="5" cols="30">Zeile 1
Zeile 2
```

Superviel Text</textarea>

Zeile 1 Zeile 2

Superviel Text

Weitere Informationen: https://developer.mozilla.org/de/docs/Web/HTML/Element/textarea

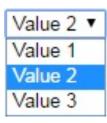
HTML Elemente Formulare



<select>...</select>

- Erlaubt die Auswahl eines Wertes (ähnlich zu Radiobuttons)
- <option>-Tags enthalten die Möglichkeiten
- Beispiel





Weitere Informationen: https://developer.mozilla.org/docs/Web/HTML/Element/select

HTML Elemente



Weitere nützliche HTML Elemente

- Seitenstrukturierung
 - > header, nav, aside, main, section, article, footer
- Textstrukturierung
 - > pre, figure, dt, dd, hr
- Textauszeichung
 - > cite, var, time, del, ins
- Multimedia
 - > canvas, svg, iframe, audio, video
- Formulare
 - > fieldset, legend, progress

Weitere Informationen: https://wiki.selfhtml.org/wiki/HTML

HTML Elemente







Dokumentstruktur	Textauszeichnung	Multimedia und Grafiken	
 html, head, body Kopfdaten 	a (Verweise)b, em, i, kbd, mark, s, small, strong, sub, sup, u	 img, picture, map, area, canvas, svg, math iframe, embed, object, param 	
title, meta, link, style, base	 cite, q (Zitate) dfn, abbr code, var, samp time ruby, rt, rp bdi, bdo 	audio, video, source, track	
Seitenstrukturierung		form, fieldset, legend, label, datalist	
 body, header, nav, aside, main, section, article, footer, address h1, h2, h3, h4, h5, h6 (Überschriften) 	 br, wbr (Zeilenumbrüche) del, ins (Änderungsmarkierungen) span 	 input, button, select, optgroup, option, textarea, keygen output, progress, meter 	
Textstrukturierung	Links (Verweise)	Skripte	
 h1, h2, h3, h4, h5, h6 (Überschriften) p, pre, blockquote, figure, figcaption ol, ul, dl, li, dt, dd (Listen) 	= a = map, area	 script, noscript, canvas content, decorator, element, shadow, template 	
• hr • div	Tabellen • table	Interaktive Elemente	
	 table caption col, colgroup thead, tbody, tfoot tr th, td 	 details, summary dialog menu, menuitem, command 	
		Besondere Attribute	
		UniversalattributeEventhandler	

Quelle: https://wiki.selfhtml.org/wiki/HTML

HTML5 – Event-Attribute



Einer der wichtigen Einsatzszenarien werden durch die Ereignis-Attribute in HTML erreicht

- Basiswerkzeug, um Dynamik in Web-Seiten einzubauen
- Fest verknüpft mit der später noch behandelten Fähigkeit, JavaScript-Code im Browser auszuführen
- Code wird bei bestimmten Handlungen ausgeführt
- Hinweis: Es gibt auch native JavaScript-Möglichkeiten
- Vielfältige Event-Attribute
 - > Fenster-Attribute sind mit dem Body-Element verknüpft (z.B. onunload)
 - > Formular-Attribute unterstützen Ereignisse in Formularen (z.B. onfocus)
 - > Tastatur-Attribute reagieren auf Tastenaktionen (z.B. onkeydown)
 - > Maus-Attribute auf Mausoperationen (z.B: onclick)
 - > Drag-Attribute sind mit Drag-Operationen der Maus verknüpft (z.B. ondrag)
 - > Clipboard-Attribute sind mit Markierungsaktionen verknüpft (z.B. oncopy)
 - > Media-Attribute auf Ereignisse bei der Behandlung von Bildern, Audio und Videos (z.B. onended)