HTML(5) - HyperText Markup Language

Dozent: Prof. Dr. Michael Eichberg

Kontakt: michael.eichberg@dhbw-mannheim.de, Raum 149B

Version: 2024-05-07



Folien: https://delors.github.io/web-javascript/folien.rst.html

https://delors.github.io/web-javascript/folien.rst.html.pdf

Fehler auf Folien melden:

https://github.com/Delors/delors.github.io/issues



1. EINFÜHRUNG

Prof. Dr. Michael Eichberg

HyperText Markup Language (HTML)

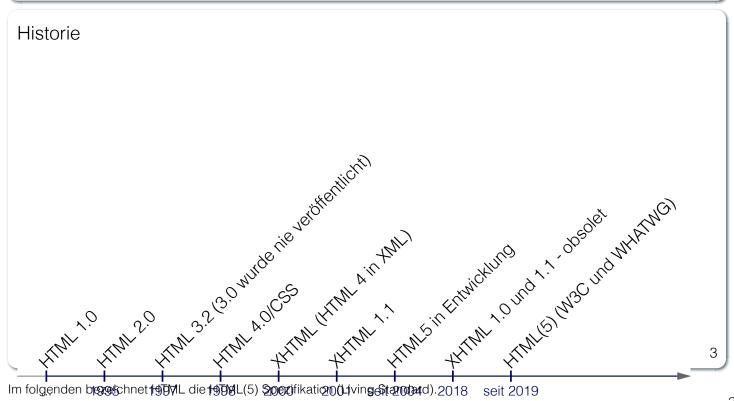
- Sprache zur Beschreibung der Darstellung von Inhalten (Markup Language), zwischen denen "navigiert" werden kann (Hypertext).
- Auszeichnungssprache abgeleitet aus SGML (Standard Generalized Markup Language).

1

Verwendungszweck:

- Webseiten
- Progressive Web-Apps
- Desktop Apps (z.B. mit Electron)

2



S

HTML vs. XML Syntax vs. DOM

Die *HTML Spezifikation* definiert eine abstrakte Sprache zur Beschreibung von Dokumenten. XML und HTML sind konkrete Syntaxbeschreibungen dieser abstrakten Sprache.

- HTML ist eine Beschreibungssprache für entsprechende Dokumente.
- XML ist eine Beschreibungssprache, die auch für HTML verwendet werden kann. (MIME Type: application/xhtml+xml) (nicht mehr empfohlen)
- Das DOM (■ Document Object Model) ist die In-Memory Darstellung eines Dokuments.

Das DOM ist ein API, um HTML Dokumente zu manipulieren.

Hinweis

Das DOM, die HTML-Syntax und die XML-Syntax können nicht alle denselben Inhalt darstellen.

1

Beispiele für Unterschiede

- Namespaces werden nicht von der HTML-Syntax unterstützt, aber sowohl vom DOM als auch der XML-Syntax unterstützt.
- **noscript** wird nur in HTML Dokumenten unterstützt.
- Kommentare, die --> enthalten werden nur vom DOM unterstützt.

HTML in a Nutshell

HTML-Dokumente bestehen aus einem Baum von Elementen und Text.

HTML Dokument

DOM

```
—DOCTYPE: html
—html lang="de"
—head
—title
— #text: Eine Webseite
— #text: ←
—body
— #text: ←
— h1
```



2. Aufbau von HTML Dokumenten

Prof. Dr. Michael Eichberg

HTML Dokumente

Die Dokumente müssen aus den folgenden Teilen in der angegebenen Reihenfolge bestehen:

- Optional ein einzelnes U+FEFF BYTE ORDER MARK (BOM) Zeichen.
- Eine beliebige Anzahl von Kommentaren und ASCII-Whitespace.
- Eine DOCTYPE Deklaration .
- Eine beliebige Anzahl von Kommentaren und ASCII-Whitespace.
- Das Dokumentenelement in Form eines html-Elementen[1].
- Eine beliebige Anzahl von Kommentaren und ASCII-Leerzeichen.

—HTML Spezifikation

[1] HTML ist nicht case-sensitive, d.h. html und HTML sind gleichwertig. Wir verwenden jedoch immer die Kleinschreibung.

Allgemeiner Aufbau von HTML Elementen

```
Start Tag | Inhalt des Elements | End Tag | Company | Co
```

Warnung

Die Spezifikation verlangt nicht in allen Fällen ein Start und/oder Endtag. Es ist jedoch eine gute Praxis, diese immer zu verwenden, wenn ein Start und/oder Endtag möglich ist.

Im Fall von Elementen ohne Endtag (z. B. <wbr/>br> darf auch keines hinzugefügt werden!

8

Beispiel

Ist ein gültiges Dokument. Es ist jedoch **keine** gute Praxis (hier wurde das *Start Tag* des **html** Elements weggelassen.

Typen von HTML Elementen

Sechs Typen von HTML-Elementen werden unterschieden:

Void elements: area, base, br, col, embed, hr, img, input, link, meta,

source, track, wbr

Das template Element:

template

Raw text elements:

script, style

Escapable raw text elements:

textarea, title

Foreign elements: Elemente aus dem MathML- und SVG-Namensraum. **Normal elements:** Alle weiteren HTML Elemente sind *normale Elemente*.

Attribute in HTML

Attribute liefern Informationen über das Element.

... Attribute

- Attribute kommen nur beim Start Tag vor.
- Attribute (in HTML) können, müssen aber kein Wert haben (Boolsche Attribute).
- Attributwerte sollten in Anführungszeichen (■ quoted) (entweder: "oder ') stehen, müssen aber nicht.

Werte ohne Anführungszeichen dürfen keine Leerzeichen oder Anführungszeichen enthalten.

Konkrete Attributwerte aus der HTML Spezifikation sind case-insensitive; andere Werte sind es nicht.

10

Im Allgemeinen sollten Attributwerte klein geschrieben werden. Selektoren in CSS und JavaScript sind casesensitive.

Z. B. ist <input type="text"> und <input type="TEXT"> gleichwertig, aber <div id="text"> und <div id="Text"> nicht!

HTML Grundgerüst

HTML <head> Element - "obligatorische" Elemente

Im Head sollten immer die folgenden Informationen deklariert werden:

Titel: Der Titel des Dokuments mit Hilfe des **title** elements

Zeichensath: Der verwendete Zeichensatz mit Hilfe des passenden meta elements:

<meta charset="utf-8">

Viewport: Konfiguration des *Viewports*[2] (insbesondere für mobile Geräte relevant):

<meta name="viewport" content="width=device-width,</pre>

initial-scale=1.0">

[2] Der Viewport (Ansichtsbereich) des Browsers ist der Bereich des Fensters, in dem der Webinhalt zu sehen ist.

Insbesondere Mobilgeräte haben oft entweder eine geringere Auflösung als Desktop-Computer oder verwenden HiDPI Screens. Beides führt dazu, dass die Webseiten nicht wie gewünscht aussehen. In diesem Fall verwenden die Browser für die Webseiten einen virtuellen Viewport mit (z. B.) 960px und skalieren dann die Seite auf 390px herunter. Wenn dieses Verhalten nicht gewünscht ist - z. B. weil die Seite Responsive ist oder von vorneherein auf mobile Endgeräte ausgerichtet ist, dann ist auf jeden Fall eine Viewport Konfiguration notwendig.

Device	Viewport Size (width x height)	Device Resolution (width x height)
iPhone 12	390 x 844	1170 x 2532
iPhone 12 Mini	360 x 780	1080 x 2340
iPhone 12 Pro	390 x 844	1170 x 2532
iPhone 12 Pro Max	428 x 926	1248 x 2778

Siehe: https://experienceleague.adobe.com/en/docs/target/using/experiences/vec/mobile-viewports.html?lang=de für weitere Details.

HTML <head> Element - weitere Elemente

Im Head können weitere Informationen und Pragmas deklariert werden bzw. sollten dort deklariert werden, wenn sie benötigt werden:

Skripte: <script [src="script.js"

[defer|async]]></script>

CSS: Mittels k rel="stylesheet" href="style.css">

oder <style>...</style>

Favorite lcon: link rel="icon" type="image/png"

href="/img/icon.png" />

Pragmas: http-equiv ≜ HTML equivalent die Informationen könnte auch im HTTP Header

stehen.

<meta http-equiv="Content-Type"</pre>

content="text/html; charset=utf-8"> (alt)

<meta http-equiv="Content-Security-Policy"</pre>

content="default-src https:" />

Äquivalente HTTP Header Definition:

Content-Security-Policy: default-src https:

Benannte Meta-Daten:

<meta name="author" content="Michael Eichberg">

13

Content Security Policies

Semantisches HTML vs. "generisches" HTML

Semantisches HTML

- Verwendung von HTML Elementen, die die Bedeutung des Inhalts klar machen.
- Bessere Zugänglichkeit
- Bessere Suchmaschinen-Optimierung

Beispiel Elemente

```
<header>, <footer>, <nav>,
<article>, <section>, <aside>,
<main>, <figure>, <address>, <b>,
<s>, ...
```

Nicht-Semantic HTML

- Verwendung von <div> und Elementen, um den Inhalt zu strukturieren.
- Keine klare Bedeutung des Inhalts.

Semantisches HTML vs. "generisches" HTML - Beispiel

```
<div>
                                                   <header>
   <span>Zwei Wörter</span>
                                                       <h1>Zwei Wörter</h1>
                                                           <a>Ein Wort</a>
       <a>Ein Wort</a>
       <a>Ein Wort</a>
                                                           <a>Ein Wort</a>
   </div>
                                                       </nav>
</div>
                                                   </header>
<div>
                                                   <main>
       <div>Viele Wörter</div>
                                                           <h1>Viele Wörter</h1>
   </div>
                                                       </header>
   <div>
                                                       <section>
       <div>Erste Worte</div>
                                                           <h2>Erste Worte</h2>
       <div>DaDaDa</div>
                                                           DaDaDa
       <div>BlaBlaBla</div>
                                                           BlaBlaBla
   </div>
                                                       </section>
</div>
                                                   </main>
<div>
                                                   <footer>
   <span>Alle Worte
                                                       Alle Worte
</div>
                                                   </footer>
```

Semantische Informationen im DOM zu haben, ist insbesondere für die Barrierefreiheit notwendig.

Alternativ zur Verwendung von semantischen Elementen können auch generische Attribute mit dem **role** Attribute versehen werden, um die Bedeutung des Elements zu spezifizieren: <div role="navigation">...</div>

Verwenden Sie HTML zur Strukturierung von Inhalten, und nicht, um das Aussehen der Inhalte zu definieren.

Das Aussehen ist Sache von CSS.

Strukturierung von Dokumenten

■ header, footer, nav, article, section, aside, main, figure, address, ...

In Hinblick auf die konkrete Semantik eines Elements gibt es Unterschiede wo und wie oft diese verwendet werden.

Ein **footer** Element innerhalb eines **article** Elements hat eine andere Bedeutung als ein **footer** Element auf oberster Ebene.

Ein main Element sollte nur einmal pro Dokument verwendet werden.

■ Überschriften: h1, h2, h3, h4, h5, h6

Attribute

Boolsche Attribute:

sind wahr, wenn diese angegeben sind und falsch andernfalls. Z. B. <input id="the-checkbox" type="checkbox" checked>.

Aufgezählte Attribute (enumerated values):

definieren eine begrenzte Anzahl von gültigen Werten sowie einen Default Wert, der verwendet wird, wenn kein Wert angegeben ist, aber das Attribut verwendet wird.

Globale Attribute: können für jedes Element verwendet werden; sind aber nicht immer sinnvoll.[3]

Globale HTML Attribute sind Z. B. id, class, data-*, autofocus, role, lang, style, popover, tabindex.

Event Handler Attribute: onclick, onclose, ...

[3] Globale Attribute

..18

Boolsche Attribute sollten in JavaScript durch hinzufügen bzw. löschen gesetzt werden (und nicht die Manipulation des Attributwertes).

```
const checkbox = document.getElementById("the-checkbox");
checkbox.removeAttribute("checked");
checkbox.setAttribute("checked");
```

Ausgewählte globale Attribute

class:

data-*:

id: verwendet, um ein Element eindeutig zu identifizieren

(Welches man in CSS oder JavaScript per Selektor referenzieren kann.) - als Ziel von Hyperlinks () - im Rahmen der Unterstützung von Barrierefreiheit - der Wert ist case-sensitive

Best Practice: Kleinbuchstaben und Bindestriche verwenden (Unterstriche sind erlaubt aber im Zusammenhang mit CSS nicht optimal).

 das class-Attribut ermöglicht es Elemente mit CSS und JavaScript anzusprechen

dient keinem anderen Zweck in HTML

wird sehr häufig von Frameworks und Bibliotheken verwendet

style: Das **style**-Attribut ermöglicht die (ad-hoc) Anwendung von Inline-Styles auf das entsprechende Element (nicht empfohlen).

Das data-*-Attribut ermöglicht es, benutzerdefinierte Daten an das Element zu binden, die von JavaScript verwendet werden können. * kann ein beliebiger Name sein, aber nicht xml oder:

of the first term of the first

HTML - logische Gruppierung von Text

Paragraphen: Inhalt

Zitate:
 <blookquote> und <q> (für kurze Zitate innerhalb eines

Absatzes)

Das Inline-Zitat-Element <q> fügt der Sprache entsprechende

Anführungszeichen hinzu.

Beispiel

"Ein Zitat" (deutsch) "A quote" (englisch)

<q lang="de">Ein Zitat</q>

<q lang="en">A quote</q>

Betonung: (■ emphasized) und

Randbemerkungen:

<small> - für Randbemerkungen und Kleingedrucktes (d. h.
small steht nicht für unwichtige(re)n Text oder die Schriftgröße)

Veraltet bzw. nicht mehr korrekt:

<S>

Zitierung: cite - für den **Titel** eines Werkes oder einer Publikation

Definitionen: <dfn [title="der definierte Begriff"]> - für die

Definition eines Begriffs

Abkürzungen: <abbr title="HyperText Markup"

Language">HTML</abbr> - für Abkürzungen

Zeitangaben: <time datetime="2021-10-01">1. Oktober

2021</time> - für Zeitangaben

Code: <code> - für Code; für das Darstellen von Code-Beispielen wird

code häufig mit kombiniert; die Sprache des Codes wird
dann über ein class Attribute spezifiziert (z. B. <code</pre>

class="language-java">...</code>)

Programmierkontexten

(Tastatur-)Eingaben:

< kbd> - für Tastatureingaben oder andere Benutzereingaben

```
Drücken Sie <kbd>cmd</kbd> + <kbd>c</kbd> zum Kopieren.
```

Hoch-/Tiefstellung:

<sup> und <sub> - für Hoch- und Tiefstellung, die nicht typographisch Zwecken dient, sondern inhaltlichen Zwecken.

H₂0 steht für Wasser.

Text mit abweichender Bedeutung:

<i>- Text, der von normaler Prosa abweicht wie z. B. eine taxonomische Bezeichnung, ein technischer Begriff, ...

Brot besteht aus <i>Mehl</i>.

Text mit erhöhter Aufmerksamkeit:

- Text, der erhöhte Aufmerksamkeit erfordert, aber nicht unbedingt betont werden muss; z. B. Schlüsselwörter in einem Artikel.

Das Wetter ist heute schön.

Text mit erhöhter Bedeutung:

<mark> - Text, der hervorgehoben werden soll, z. B.

20

Es gibt weitere Elemente, die für spezielle Anwendungsfälle verwendet werden können. Siehe WHATWG.

HTML Text - physische Auszeichnung von Text

Vorformatierter text:

< - color="block"> - für Text, der so angezeigt werden soll, wie er

geschrieben wurde)

Zeilenumbrüche:
 - für Zeilenumbrüche, die inhärenter Teil der Daten sind wie

zum Beispiel bei Adressen. D. h. sollte nicht innerhalb von Text

verwendet werden!

Optionale Zeilenumbrüche:

<wbr> (■ word break opportunity`) - ein optionaler

Zeilenumbruch

(Beispiel: Er schrie: <q

lang="de">Lasst<wbr>Mich
In<wbr>Ruhe!)

HTML - und div

- und <div> sind generische Container-Elemente, die verwendet werden, um Text oder andere Elemente zu gruppieren.
- ist ein Inline-Element
- <div> ist ein Block-Element
- beide werden häufig verwendet, um CSS-Klassen zuzuweisen, um den Inhalt zu gruppieren oder um den Inhalt zu manipulieren.

HTML - data

- Das data-Attribut ermöglicht es, benutzerdefinierte Daten an das Element zu binden, die von JavaScript verwendet werden können.
- Die Daten werden im **value** Attribut gespeichert.

<data value="8">Acht</data>

HTML - Links

- Hyperlinks werden mit dem <a> Element erstellt.
- Der href-Attribut enthält die Adresse des Ziels (innerhalb des gleichen Dokuments, auf einer anderen Webseite, per E-Mail, ...)

```
1 <a href="https://www.dhbw-mannheim.de">DHBW Mannheim</a>
2 <a href="#teachers">Unsere Lehrenden</a>
3 <a href="mailto:michael.eichberg@dhbw-mannheim.de">Email: Michael Eichberg</a>
```

- 1. Externer Link
- 2. Interner Link (I link fragment identifier) auf ein Element mit der ID teachers
- 3. E-Mail Link kann ergänzt werden durch subject und body Parameter innerhalb des href Attributs.
- Das target-Attribut ermöglicht die Definition des Browsing-Kontextes für die Link-Navigation (und die Formularübermittlung).

```
1 <a target="_blank" href="https://www.dhbw-mannheim.de">DHBW Mannheim</a>
2 <a target="_self" href="https://www.dhbw-mannheim.de">DHBW Mannheim</a>
3 <a target="_top" href="https://www.dhbw-mannheim.de">DHBW Mannheim</a>
4 <a target="_parent" href="https://www.dhbw-mannheim.de">DHBW Mannheim</a>
5 <a target="dhbw" href="https://www.dhbw-mannheim.de">DHBW Mannheim</a>
```

- 1. Öffnet den Link in einem neuen Fenster oder Tab
- 2. Öffnet den Link im gleichen Browsing-Kontext
- 3. Öffnet den Link im obersten Browsing-Kontext
- 4. Öffnet den Link im übergeordneten Browsing-Kontext
- 5. Öffnet den Link im Browsing-Kontext mit dem Namen dhbw (Beispiel: DHBW Mannheim)

_self, _top und _parent sind relativ zum aktuellen Browsing-Kontext und unterscheiden sich nur, wenn die Seite in einem Frame oder einem iframe angezeigt wird.

Das rel-Attribut legt die Art des Links fest und definiert die Beziehung zwischen dem aktuellen Dokument und der Ressource, auf die der Hyperlink verweist. (Z. B. rel="license", rel="author" oder rel="noopener"; siehe MDN rel attribute)

24

Durch die Zuweisung zu einem Browsing-Kontext kann verhindert werden, dass die selbe Seite X mal geöffnet wird, wenn ein Nutzer auf den Link klickt.

HTML - Lists

Drei Arten von Listen werden unterstützt, die beliebig verschachtelt werden können:

1. geordnete Listen:]>

```
start="0">
                                   0. Erster Punkt
 Erster Punkt
                                   1. Zweiter Punkt
 Zweiter Punkt
                                  10. Dritter Punkt
 value="10">Dritter Punkt
</01>
```

2. ungeordnete Listen:

```
ul>
                              Erster Punkt
 Erster Punkt
                              Zweiter Punkt
 Zweiter Punkt
                              Dritter Punkt
 Dritter Punkt
```

3. Definitionslisten: <dl>

```
<dl>
                                           Erster Begriff
                                                            Erklärung des 1. Begriffs
  <dt>Erster Begriff</dt>
                                           Zweiter Begriff
                                                            Erklärung des 2. Begriffs
  <dd>Erklärung des 1. Begriffs</dd>
  <dt>Zweiter Begriff</dt>
  <dd>Erklärung des 2. Begriffs</dd>
                                                                                        25
```

HTML - Navigation

- Das <nav> Element wird verwendet, um Navigationslinks zu gruppieren.
- Insbesondere für Screenreader und die Suchmaschine relevant.

HTML - Tabellen

Verwendet für die Darstellung von tabellarischen Daten mit Zeilen und Spalten.

Hinweis

Die Verwendung von sollte sich nach dem Inhalt richten!

Tabellen sollten nicht zum Layout von Webseiten verwendet werden.

Aufbau von Tabellen:

```
Logische
 <caption>Logische Operation/caption>
                          Operation
   not xor1
                             1
                               0
                          xor
 </thead>
                             1
                          1
                               0
   11
                          0
                             0
                               1
   001
 <tfoot></tfoot>
```

- Zellen, die über mehrere Spalten oder Zeilen gehen können mit Hilfe von **colspan** und **rowspan** Attributen definiert werden.
- Spalten und Zeilen können mit Hilfe von **<col>** und **<colgroup>** Elementen definiert werden.

HTML - Images

Bilder werden mit dem Element eingebunden.

```
<img src="path/filename" alt="descriptive text" />
```

- Das src-Attribut enthält die Adresse des Bildes.
- Das alt-Attribut enthält eine Beschreibung des Bildes, die angezeigt wird, wenn das Bild nicht geladen werden kann.
- Das width und height-Attribut können und sollten verwendet werden, um die Größe des Bildes festzulegen.
- Lazy loading ist durch die Verwendung des loading Attributs möglich (loading="lazy").
- Folgende Bildformate werden breit unterstützt: jpg, png, gif, svg und webp.
- Responsive Images werden über das srcset Attribut unterstützt:

```
<img src="images/dhbw.png" alt="Logo der DHBW"</pre>
    srcset="images/dhbw.png 400w, images/dhbw-xl.jpg 800w"
    sizes="(max-width: 800px) 400px, 800px" />
```

Weitere Responsive Features) werden mittels CSS ermöglicht. Um zum Beispiel zu verhindern, dass ein Bild größer als eine Textzeile wird, kann folgendes CSS definiert werden:

```
img {
    max-inline-size: 100%;
    block-size: auto;
```

HTML - Formulare

Formulare werden mit dem <form> Element erstellt.

- **action** enthält die Adresse, an die die Formulardaten gesendet werden.
- method definiert die Methode, die zum Senden der Daten verwendet wird (GET oder POST).
- name setzt den Namen des Formulars.
- target enthält den Namen des Browsing-Kontexts, in dem die Antwort angezeigt wird.
- autocomplete ermöglicht das automatische Ausfüllen von Formularen.
- novalidate verhindert die Validierung der Formulardaten durch den Browser.
- accept-charset definiert die Zeichencodierung, die zum Senden der Formulardaten verwendet wird.

Formularelemente werden mit dem <input> Element erstellt.

- type definiert den Typ des Formularelements.
- name definiert den Namen des Formularelements.
- value definiert den Wert des Formularelements.
- placeholder definiert den Platzhaltertext des Formularelements.
- required definiert, ob das Formularelement erforderlich ist.
- disabled definiert, ob das Formularelement deaktiviert ist.
- autofocus definiert, ob das Formularelement den Fokus erhält.

Beispiel

```
<form method="GET"
    name="Folienauswahl">
    <label for="slide">Folie:</label>
    <select name="ld-slide-no" id="slide">
        <option value="8">Elemente</option>
        <option value="10">Attribute</option>
        <option value="29">Formulare</option>
        </select>
        <input type="submit" value="Submit">
        </form>
```

Für weitere Informationen bzgl. Formulare siehe MDN Web Docs oder Web.dev.

HTML - Zusammenfassungen und Details

HTML unterstützt verschiedene interaktive Elemente:

■ Anzeigen von optionalen Details mit Hilfe des <details> Elements.

HTML - Dialoge

■ Dialoge werden mit Hilfe des <dialog> Elements erstellt. Dialoge sind spezielle Fenster, die den Fokus auf sich ziehen und die Interaktion mit dem Rest der Seite unterbrechen - falls diese modal sind.

```
Beispiel[4]:

<dialog>
    <h1>Dialog</h1>
    Dialog Inhalt
    <button formmethod="dialog">
        OK
      </button>
      <button autofocus>Abbrechen</button>
</dialog>
Open Dialog

Open Dialog
```

[4] JavaScript Code zum Öffnen des Dialogs wird hier nicht gezeigt.

HTML Entities

Ausgewählte Zeichen können (in manchen Kontexten) nur durch HTML Entities dargestellt werden:

```
< durch &lt; oder &#60; (<)</pre>
```

- > durch: > oder > (>)
- & durch: & oder & (&)
- " durch: " oder " (")
- _ durch: ()

Benannte Zeichen

HTML Erweiterbarkeit

- Hinzufügen von Meta-daten (<meta name="" content="">)
- class Attribute
- "Custom Elements" (z.B. <my-element>)
- Autoren können APIs mit Hilfe des JavaScript-Prototyping-Mechanismus erweitern

Referenzen

- MDN Web Docs
- acanisuse.com: Unterstützung von HTML, CSS etc. Features
- HTML (Living Standard) (aka HTML5)
- HTML DOM