## Chapter 9: HTML/XHTML I

## **References:**

- Erik Wilde: World Wide Web Technische Grundlagen. Springer, 1999, ISBN 3-540-64700-7, 641 Seiten.
- Eric Ladd, Jim O'Donnell, et al.: Using HTML 4, XML, and Java 1.2, Platinum Edition. QUE, 1999, ISBN 0-7897-1759-X, 1282 pages.
- Rainer Klute: Das World Wide Web. Addison-Wesley, 1996, ISBN: 389319763X.
- Dave Raggett, W3C: HTML 3.2 Reference Specification. [http://www.w3.org/TR/REC-html32.html]
- Dave Raggett, Arnaud Le Hors, Ian Jacobs (Eds.): HTML 4.01 Specification. W3C, Dec 24, 1999. [http://www.w3.org/TR/html4/]
- User's Guide to ISO/IEC 15445:2000 HyperText Markup Language (HTML) [http://www.cs.tcd.ie/15445/UG.html]
- XHTML [tm] 1.0: The Extensible HyperText Markup Language. W3C, Jan 26, 2000. [http://www.w3.org/TR/xhtml1]
- Stefan Münz: HTML-Dateien selbst erstellen SELFHTML. [http://www.netzwelt.com/selfhtml/] [http://www.teamone.de/selfaktuell/]
- Ian Graham: Introduction to HTML. [http://www.utoronto.ca/webdocs/HTMLdocs/NewHTML/htmlindex.html]
- NCSA Beginner's Guide to HTML (no longer maintained). [http://www.ncsa.uiuc.edu/General/Internet/WWW/]

## Objectives

After completing this chapter, you should be able to:

- develop web pages in strict HTML 4.0/XHTML 1.0.
- write syntactically correct HTML/XHTML.
- read the HTML and XHTML specifications.
- evaluate whether something is possible in HTML.









also to a destination anchor inside it (see below).





Hypertext-Links: a (6)

- What happens when the user activates a link (clicks on it), depends on the URI, the media type of the referenced resource, and the browser configuration:
  - HTML documents, ASCII text files, images, etc. are normally displayed in the same browser window (they replace the current document).
  - ♦ The browser can be configured in such a way that it calls helper applications to display certain media types (e.g. ghostview for postscript).
  - $\diamond\,$  It is also possible to store the data in a file.













Hypertext-Links: a (13)

ATTLIST</th <th>a</th> <th>%attrs;</th> <th></th> <th></th>	a	%attrs;		
		charset	%Charset;	#IMPLIED
		type	%ContentType;	#IMPLIED
		name	CDATA	#IMPLIED
		href	%URI;	#IMPLIED
		hreflang	%LanguageCode;	#IMPLIED
		rel	%LinkTypes;	#IMPLIED
		rev	%LinkTypes;	#IMPLIED
		accesskey	%Character;	#IMPLIED
		shape	%Shape;	rect
		coords	%Coords;	#IMPLIED
		tabindex	NUMBER	#IMPLIED
		onfocus	%Script;	#IMPLIED
		onblur	%Script;	#IMPLIED>







attribute name is declared as CDATA (arbitrary string).









## The Base URI (2)

- E.g. when a HTML document is sent via email, the document itself has no URI, so that if one should use relative URIs in this case, "base" is required.
- Defining a base URI permits to move or copy the document without moving the documents that are referenced via relative URIs within it.

However, even if one copies the entire set of documents, relative URIs will then still refer to the old place defined with "base". In this case, it would have been better not to use "base".





9 - 26



use the information contained in these links.



fetch it when the user wants to print the document.





The HTML 4.01 specification mentions the following kinds of links: Alternate, Stylesheet, Start, Next, Prev, Contents, Index, Glossary, Copyright, Chapter, Section, Subsection, Appendix, Help, Bookmark.

The HTML 3.2 specification mentions: top, contents, index, glossary, copyright, next, previous, help, search. In the DTD for HTML 4.01 and XHTML, a comment mentiones start, contents, previous, next, index, end, help, stylesheet, script, alternate. The HTML 4.01 specification states that capitalization is not relevant, but probably one should use all lowercase for XHTML.

It was once recommended to put the author's email address in this way into the document: <link rev="made" href="mailto:brass@acm.org" />
















Images: img(4)• In HTML 3.2, alt was optional. In HTML 4, it is required. For pictures that do not contain information and are only intended to make the web page more beautiful, the HTML specification recommends alt="". • In HTML 4, img has in addition an optional attribute longdesc, which contains the URI of a longer description of the picture. E.g. a blind person can select this if the alt-text sounds interesting or does not contain enough information.









- With the above settings for align, the picture is simply treated like a big letter.
- However, with align="left" or align="right", one can move the picture to the left or right side of the window and let the text "flow around it":



• In HTML 3.2/HTML 4 Transitional,

```
<br clear="left" />
```

- can be used to move the writing position below the picture on the left side.
- In HTML 4 Strict, img has no attribute align, and br has no attribute clear. One must use stylesheets instead.
- Often, pictures are also used together with tables.



- Furthermore, img had the following attributes in HTML 3.2/HTML 4 Transitional:
  - hspace: Horizontal empty space to the left and right of the picture.
  - vspace: Vertical empty space above and below the picture.
  - border: Size of the border that the browser draws around clickable images.

E.g. with border="0", this border is not displayed.

Images: img (11)

 In HTML 4 (Strict and Transitional), img has an attribute name, in order to reference the picture in stylesheets and scripts (programs). However, one should use id instead.

The attribute name is kept only for compatibility reasons (although it was never officially part of HTML 3.2). It cannot be used for destination anchors (at least, it is not intended for that purpose).

 Finally, there are two further attributes: usemap and ismap (in HTML 3.2, HTML 4 Strict+Transitional).
They are used for "image maps" (see below).





- "Image maps" make it possible that the user reaches different documents depending on the point or region in the picture where he/she clicks.
- There are two kinds of image maps:
  - Server-Side Image Maps: The browser tells the server the mouse coordinates. The server delivers different data based on these coordinates.
  - ♦ Client-Side Image Maps: The HTML file contains a list of regions in the picture, each with its own URI. The browser loads the selected URI.



"ismap" is the only possible value of the attribute (besides a "NIL" value when the attribute is not specified).







• Within the map-element, area-elements are used to specify the regions and their associated URLs.







TTIME 4. area-clements are ignored when the map contain



Clent-Sid	e Imag	ge Map	5 (6)
• The element ty	pe area is (	declared as f	ollows:
ELEMENT area</td <td>- 0</td> <td>EMPTY&gt;</td> <td></td>	- 0	EMPTY>	
ATTLIST area</td <td>%attrs:</td> <td></td> <td></td>	%attrs:		
	shape	%Shape:	rect
	coords	%Coords:	#IMPLIED
	href	%URI:	#IMPLIED
	nohref	(nohref)	#IMPLIED
	alt	%Text;	#REQUIRED
	tabindex	NUMBER	#IMPLIED
	accesskey	%Character;	#IMPLIED
	onfocus	%Script;	#IMPLIED
	onblur	%Script;	#IMPLIED>









Typically inside a pre element. The problem is to align the second and following columns at a common edge.





 Because CSS1 stylesheets do not support tables, the HTML 4 table elements still have presentationoriented attributes.

With CSS2 these attributes become unnecessary.













Universität Halle, 2011














because the first two columns are covered by the large entry in the row above.



• The caption-text is displayed above the table. In HTML 3.2/HTML 4.01 Transional: also align=bottom possible. 9-77











- Note that tfoot must be specified before tbody, although the footer is printed at the bottom of the table.
  - In this way, it might be possible to print the first page of a long table while data still arrive.
- It is possible to use several tbody-elements in order to partition the data rows into several groups.

E.g. one can request separating lines between groups of rows, see below.











9-87



- If col/colgroup are used, no table row may have more entries than the declared number of clumns.
- Note that col/colgroup do not produce any output.
  Column headings must still be specified with th.





were added at the end.





 $\diamond$  box, border: All four sides.





or one defines in each td-element with headers a list of id's of the relevant th-elements.





```
<TH scope=col abbr="">Prozent
```



















Most frameset-elements contain only one of the two attributes. However, it is also possible to use both in order to create a grid. It is said that this might confuse old browsers.





the last frame 200.



• The main file never contains the frame contents.
```
Frame Definition (5)
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN"
       "http://www.w3.org/TR/html4/frameset.dtd">
<html>
<head><title>Example for Frames</title></head>
<frameset rows="70,*">
   <frame src="f_banner.gif" />
   <frameset cols="20%,80%">
       <frame src="f_toc.html" />
       <frame src="f_doc1.html" name="doc" />
   </frameset>
<noframes><P>Your browser does not understand frames. Sorry.
         <!-- Bad Style. Better display list of links -->
</noframes>
</frameset>
</html>
```







In HTML, one writes <frame ... noresize>. In XHTML, one must write <frame ... noresize="noresize" />.



- Normally, the browser uses automatically scrollbars, if the contents does not fit into the frame.
- With scrolling=no the scrollbars can be switched off, with scrolling=yes they are activated even if the contents does fit.

The default is scrolling=auto.

• Switching off scrollbars is bad style. How should the user see the rest of the frame contents?



## Frame Definition (11)

- title can contain a short description of the frame, and longdesc can point to a loger description.
- This is especially important if src points directly to an image file.

In contrast to the img-element, there is otherwise no alternative text for people who cannot see.











