

HTML

HTML

Hypertext Markup Language Portable Content

Represent textual and image content -- not exact appearance. Better than plain ASCII, not as rich in appearance as Postscript or PDF.

Encode text and image "page" content in a way that can be displayed on a wide variety of platforms

Page is made of text content, some layout information, images, and URL links

Unsolved

That may sound like a modest goal, but it hadn't really been solved up to that time (this is part of the "Markets vs. Standards" theme)

Different Platforms

Different OSes, different window/screen sizes, different installed fonts that render at different sizes.

Users who have particular preferences on what font to display with.

References

HTML Primer

<http://www.ncsa.uiuc.edu/General/Internet/WWW/HTMLPrimer.html>

W3C

World Wide Web Consortium [W3C] -- the non-profit which tries to maintain a portable definition of HTML and other Internet related standards...

<http://www.w3.org/pub/WWW/MarkUp/>

Setting up a page in leland space

<http://www.stanford.edu/leland/howto.shtml>

View Source

To learn how a page is written, use the "View Source" option in your browser to look at the HTML source code. This is a quick way to learn.

Basic Document Structure

Text

An HTML file is just a text file. The HTTP server sends it over to the client for display in the client browser.

Structure vs. Appearance

HTML defines the basic structure of a document, but the exact appearance is determined by the browser dynamically.

Text and tags.

Tags enclosed in <>'s. Some tags occur by themselves e.g.
, others are paired to enclose some text <h1>...</h1>. Upper and lower case tags act the same.

The HTML itself is plain text — no bold, no fonts, no rulers, etc. Just a sequence of characters. You can edit the HTML text in a word-processor or text editor. Just be sure to save it as "plain text" and not some other format.

Whitespace

"Whitespace" (spaces, tabs, end-of-lines, ...) is largely ignored.

What Defines Appearance?

Appearance in the browser is determined by..

What fonts and sizes the user has chosen in their browser preferences.

The current width of their browser window.

Pour

The HTML "pours into" this structure set by the user.

Try changing the width of your window — the content defined by the HTML should rearrange to present itself in the window.

Smart HTML authors accept that the exact **appearance** is controlled by the browser, and just write HTML to define the **structure** they want.

The content will look reasonable on many different platforms -- different browsers, and even on things like Palm Pilots (essentially just a very narrow window).

See <http://cslibrary.stanford.edu/104/> for an example of a complex page which looks good on all platforms.

Basic Tags

<html>...</html>

Encloses the entire HTML source.

<head>...</head>

Encloses the "header" section of an HTML document. The header section most often just contains the document title, but it can contain other meta information about the document. Such meta information may be better

developed in the future. Currently, the Alta Vista meta content tags which aid the indexing of the document go in the header section. See the Alta Vista help page.

<title>...</title>

The title phrase for this page. Polite to always have one. Shows up as the title of the browser window, although that's a little to subtle to be useful. Also used by search engines, bookmarks, etc. as a one phrase name for this page. I frequently have use the same phrase in the <title> and a top of page <h1>.

<body>...</body>

The body is basically all of the document but the header.

<h1>Large Header</h1>

<h2>Smaller Header</h2>

<p>

Marks the beginning of a block of text which should stick together as single paragraphs. In the browser, each paragraph will be set off from its surroundings by some vertical whitespace.

All of the text after the <p> until another paragraph or header marker will be gathered up into a single paragraph. Blank lines etc. do not terminate the paragraph. The </p> is not required, but it may be someday.

..

Unordered list. Each item (below) shows up with a bullet. There's an ordered list .. variant where the items are numbered.

Each item in the list begins with an . The items themselves behave pretty much like little paragraphs. There's no tag, the beginning of one list marks the end of the previous. On the planet of HTML authors, most of the punch lines to their jokes involve looking around for the tag.

Miscellaneous Adornment Tags

**
**

Force a line break (aka a "return") in the appearance of the text. Does not introduce extra vertical whitespace...Use <p> if you want a line break and vertical whitespace.

<hr>

Horizontal Rule. Draws a horizontal line. Good for providing visual separation of logically separate elements.

Physical Styles

Indicate character appearance. These have won out over logical styles mostly

`...`

Bold.

`<i>...</i>`

Italic.

`<blink>...</blink>`

Regarded as quite-vulgar. One problem is: how do you print it on paper? I heard that Marc Andreesson admitted to being drunk when he added this feature to Mosaic.

Logical Styles

Define the logical role of the text. Also causes the text to have a distinguishing appearance in the browser. The exact appearance is up to the browser.

`...`

Emphasis. Makes the text stand out in some way. (Often italic)

`...`

Strong emphasis. Make the text stand out more so. (Often bold)

`<cite>...</cite>`

Encloses the name of something, book, movie, etc. Someday, this information may be used to build more intelligent cross-references of the web.

`<address>...</address>`

A postal or email address or reference. Typically shows up on its own line.

Special Characters

`<` -- "<"

`&` -- "&"

Example

```

<html>
<head>
<title> HTML Basic </title>
</head>

<body bgcolor=white>

<!--
  This is what a comment looks like
-->

<h1> HTML Basic </h1>

<h2> Slightly Smaller Header </h2>

<p> Here's some body text. HTML runs all this text together,
basically ignoring spaces, tabs, and blank lines until it gets to
a tag which

tells it to
be something other
than a

body text paragraph.

<p> Another body text paragraph. Each of these is typically separated
from whatever precedes it by a little vertical whitespace.

<h3>&lt;br&gt;</h3>

<p>Use the &lt;br&gt;<br>
tag to force a linebreak. Try not to use this -- it tends to produce less portable
results
since it stops accounting for the window width in use.

<h3>&lt;hr&gt;</h3>

Use &lt;hr&gt;<hr>
To introduce a horizontal line to divide sections.

<h2>Lists!</h2>

<p>
And now some exciting foods...
<ul>
<li>Bannana -- a yellow fruit</li>
<li>Apple -- a red fruit
<li>Potato -- a starchy tuber
</ul>

<h2>Physical Character Styles</h2>
<ul>

```

```

<li><b>Bold</b>
<li><i>Italic</i>
<li><tt>Fixed Width</tt>
</ul>

```

<h2>Logical Character Styles</h2>

```

<ul>
<li><cite>Cite</cite>
<li><address>Italic</address>
</ul>

```

<h2>Tables</h2>

```

<table border=1>
<tr>
  <th>Fruit</th>
  <th>Feature</th>
</tr>

<tr>
  <td>Bannana</td><td>Squishy yellowness</td>
</tr>

<tr>
  <td>Apricot</td><td>Dried out orangey-ness</td>
</tr>

<tr>
  <td>Mystery Fruit with a long name</td>
  <td>Mysteriousness -- note how this
    left aligns with the earlier features. Tables are the portable way
    to get things to line up.</td>
</tr>
</td>

</table>

</body>
</html>

```

The screenshot shows the Netscape browser window titled "Netscape: HTML Basic". The address bar contains the file path: `file:///NULL/Classes/193i%2099-3/HTML%20demo/basic.html`. The browser toolbar includes icons for Back, Forward, Reload, Home, Search, Netscape, Images, Print, Security, Shop, and Stop. Below the toolbar are navigation links: WebMail, Contact, People, Yellow Pages, Download, and Find Sites.

HTML Basic

Slightly Smaller Header

Here's some body text. HTML runs all this text together, basically ignoring spaces, tabs, and blank lines until it gets to a tag which tells it to be something other than a body text paragraph.

Another body text paragraph. Each of these is typically separated from whatever precedes it by a little vertical whitespace.

**
**

Use the `
` tag to force a linebreak. Try not to use this -- it tends to produce less portable results since it stops accounting for the window width in use.

<hr>

Use `<hr>`

To introduce a horizontal line to divide sections.

Lists!

And now some exciting foods...

- Bannana -- a yellow fruit
- Apple -- a red fruit
- Potato -- a starchy tuber

Physical Character Styles

- **Bold**
- *Italic*
- Fixed Width

Logical Character Styles

- *Cite*
- *Italic*

Tables

Fruit	Feature
Bannana	Squishy yellowness
Apricot	Dried out orangey-ness
Mystery Fruit with a long	Mysteriousness -- note how this left aligns with the earlier features. Tables are