

# *Notes from Well House Consultants*

*These notes are written by Well House Consultants and distributed under their Open Training Notes License. If a copy of this license is not supplied at the end of these notes, please visit*

*<http://www.wellho.net/net/whcotnl.html>  
for details.*

## 1.1 Well House Consultants

Well House Consultants provides niche training, primarily but not exclusively in Open Source programming languages. We offer public courses at our training centre and private courses at your offices. We also make some of our training notes available under our "Open Training Notes" license, such as we're doing in this document here.

## 1.2 Open Training Notes License

With an "Open Training Notes License", for which we make no charge, you're allowed to print, use and distribute these notes provided that you retain the complete and unaltered license agreement with them, including our copyright statement. This means that you can learn from the notes, and have others learn from them too.

You are NOT allowed to charge (directly or indirectly) for the copying or distribution of these notes, nor are you allowed to charge for presentations making any use of them.

## 1.3 Courses presented by the author

If you would like us to attend a course (Java, Perl, Python, PHP, Tcl/Tk, MySQL or Linux) presented by the author of these notes, please see our public course schedule at

<http://www.wellho.net/course/index.html>

If you have a group of 4 or more trainees who require the same course at the same time, it will cost you less to have us run a private course for you. Please visit our onsite training page at

<http://www.wellho.net/course/otc.html>

which will give you details and costing information

## 1.4 Contact Details

Well House Consultants may be found online at

<http://www.wellho.net>

[graham@wellho.net](mailto:graham@wellho.net)

technical contact

[lisa@wellho.net](mailto:lisa@wellho.net)

administration contact

Our full postal address is

404 The Spa  
Melksham  
Wiltshire  
UK SN12 6QL

Phone +44 (0) 1225 708225

Fax +44 (0) 1225 707126

# HTML – An Overview

For trainees who will be writing executable web content in any programming language, or deploying applications written by others, an elementary understanding of HTML is vital. This module provides a quick reminder of the fundamentals of HTML.

|   |   |
|---|---|
| <i>Tags</i> .....                             | 4 |
| <i>Structure of a page</i> .....              | 4 |
| <i>Special characters and new lines</i> ..... | 4 |
| <i>Some common tags</i> .....                 | 5 |
| <i>Lists, tables, etc.</i> .....              | 6 |
| <i>Which HTML standard?</i> .....             | 6 |
| <i>Validating your page</i> .....             | 7 |

HTML (HyperText Markup Language) is the text formatting language that's used to describe the content of web pages. Whilst your browser can read and display a simple text file, most web page authors prefer to use HTML which lets them suggest some formatting, supply links to other pages, etc.

Although you may be familiar with products such as FrontPage and Dreamweaver to generate HTML, you'll need to have a basic understanding of the structure when coding in other open source languages.

## 2.1 Tags

HTML directives are written in tags, between `<` and `>` signs. The first word after the `<` character is the type of tag and may be followed by a series of parameter = value pairs before the `>` sign.

Many tags mark the start of an area, and a matching tag with a `/` character before the type name marks the end of the area. Thus

```
<h1 align=center>A Centred title</h1>
```

Tag and parameter names are not case-sensitive in basic HTML, but some of the parameter values are. You are encouraged to write all tag and parameter names in lower case to ensure that they also conform to XHTML standards.

## 2.2 Structure of a page

Web pages comprise a block of HTML, usually written between tags:<sup>1</sup>

```
<html> and </html>
```

Within the HTML, the page is split into two parts.

- The Head (**<head>** to **</head>**) which contains information that does not refer to the main display of the browser; for example, details of the page author, information for search engines and probably a title that will be used for bookmarking and labelling the displayed window.

That Title will be between **<title>** and **</title>** tags.

- The body (**<body>** to **</body>**) which contains the text to be displayed in the main browser window and other information about how it is to be displayed and what else (e.g. images) is to be displayed as well. Colours for background, text and links can be specified in the **<body>** tag.

It is also recommended that you precede your web page with a document type declaration. The loosest you can use (which says that your page may include tags and attributes which are no longer encouraged) is:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"  
"http://www.w3.org/TR/html4/loose.dtd">
```

and you can replace the loose with strict if you don't want to accept deprecated tags.

## 2.3 Special characters and new lines

If you include a `"<"` sign in your text, the browser will think that a tag is coming. In order to prevent this, you can specify

```
&lt;
```

instead. You must also use

```
&amp;
```

if you want an ampersand ("`&`") in your text, and you can use a whole range of other specials too, such as

```
&pound;        £  
&copy;         ©
```

You'll be writing programs later on this course which may include any text in their output, and you'll need to filter your output to make these substitutions.

Within HTML, any (and multiple) white space characters are replaced by single

---

<sup>1</sup> You can omit these tags and the browser will usually be OK about it.

spaces, and lines are re-folded as necessary. If you want to force a line break, you should use a

```
<br>
```

tag, and if you want to leave a gap to a new paragraph you should use text between

```
<p> and </p>
```

Don't eliminate new line characters completely, though; a few `\ns` will be indispensable when you come to edit or read the text of your page.

## 2.4 Some common tags

If you're looking for a "quick and dirty" display of a table of text in a fixed width font, precede the text with `<pre>` and end it with `</pre>`. This is the preformatted tag and within the block, the compression of spaces is suppressed. It's a very useful way of keeping data in columns, but you still need to filter "<" and "&" characters!

An `<hr>` tag gives you a horizontal ruling.

An `<h1>` and `</h1>` pair lets you put in a headline size 1 (largest). You can also use `<h2>` through `<h6>` in a similar way. `<em>` and `</em>` can be used to mark a part of your text as emphasised. There are various other things one can do and this information should only be considered as very basic HTML.

Anything written between `<center>` and `</center>` tags will be centered rather than the default left-justified in your browser's window.

If you want to specify a user-selectable link to another page, use an anchor tag:

```
Go to <a href=previous.html>last newsletter</a>.
```

The `href` parameter can be any URL definition you choose – an `.html` file in the current directory in this example – but it could call another directory, another site, or even refer to a different communication protocol:

```
Download <a href=ftp://ftp.wellho.co.uk/pub/diary.txt>calendar</a>
```

Here's a page with some of these tags in use:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
<title>An HTML Reminder</title>
</head>
<body bgcolor=white text=blae
<center>
<h3>A Reminder</h3>
</center>
This page is a reminder of some of the tags that can be
included in a web page.
<p>
You may select from
<ul>
<li><a href=index.html>The home page</a>
<li>Information about
<a href=http://www.wellho.net>Well House Consultants</a>
</ul>
<hr>
<font color=green>
Graham Ellis<br>
graham@wellho.net
</font>
</body>
</html>
```

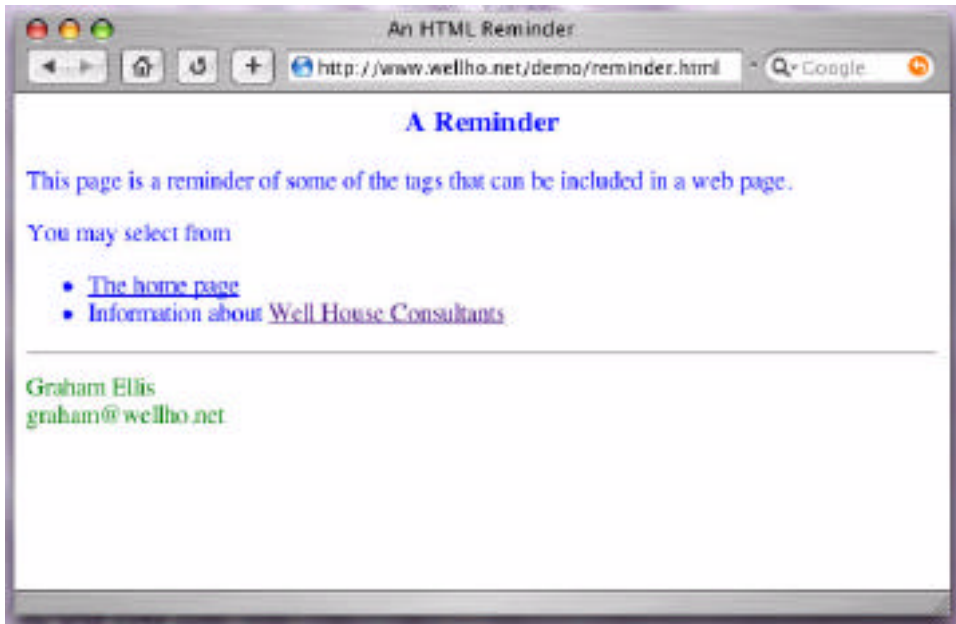


Figure 1 The HTML above, browsed using Safari on a Mac

## 2.5 Lists, tables, etc.

Within a page of HTML, you can specify an ordered list (`<ol>`) or an unordered list (`<ul>`), ending with `</ol>` and `</ul>` respectively. Individual list items should be prefixed with `<li>`.

To give better control over formatting than a list, you may want to arrange your data into a table. Tables are enclosed in `<table>` and `</table>` tags; within tables, rows are enclosed in `<tr>` and `</tr>` tags, and within each row, data elements are enclosed in `<td>` and `</td>` tags. Many parameters can be specified to tables to control their looks.

## 2.6 Which HTML standard?

The current HTML standard is HTML 4.01.<sup>1</sup> These days, the majority of visitors to your website will use the Internet Explorer browser from Microsoft, but it's as well to stick to the standard rather than use their extensions which map and prevent Netscape, Mozilla, Safari and Konqueror users from viewing your site correctly.

Remember too that users may not have a Java plugin, may change the font style and background and resize their windows at will, and can even choose not to show graphics.

Which, when all put together, means that any tags you use within a web page can be considered to be just hints to the browser which may or may not be acted on!

You should be very careful to consider your target audience for HTML, whether you're writing the HTML directly, having it generated by a software package such as DreamWeaver, FrameMaker or FrontPage, or by a Perl or PHP program.

Suggestions:

- Stick to the the recommended standard by WC3 (the World Wide Web consortium) <http://www.w3.org/>
- Do not use Microsoft extensions
- Do not rely on images to convey vital content
- If you must use Frames, also provide an alternative
- If you must use Applets, also provide an alternative
- Check your pages on Linux, Macintosh and Windows platforms

<sup>1</sup> it has been since 1999

- Check your pages on a number of different browsers
- See how your page looks when scaled right up or down

And don't assume that your users will have some special plugin, or a high resolution screen, or a fast line to the internet just because you do ;-)

In the majority of applications, users will be looking for professionally presented content and not magic bells and whistles. Even restricting yourself as described above, you should have plenty of scope to impress. Of course, if you're going to employ someone with the skills to present an effective web site, they'll need Graphic Design, Library Science, Journalism, Time and motion, marketing and computer science skills as well as a thorough understanding of the business being represented on the site. Subject for a different course!

## 2.7 Validating your page

The HTML you write will very often render correctly in most browsers even if you break a few of the rules and standards. In some ways this is great news, but how do you check your pages if you really want them to be right?

W3.org offers a free validation service at [validate.w3.org](http://validate.w3.org) where you can enter a web page URL, or a text file containing HTML, to upload and validate. Here's a sample form to test the example used in this module:

Figure 2 Validation form at w3.org

The screenshot shows the W3C Markup Validation Service interface. The browser address bar shows <http://validator.w3.org/fix-upload.html>. The page title is "W3C Markup Validation Service v0.6.6". The navigation menu includes Home, About..., News, Docs, Help & FAQ, and Feedback. Below the menu is a "Link Checker" section with the text: "This form allows you to upload files from your computer and have them validated." The main heading is "Validate Uploaded File". Underneath, there is a "Validate by File Upload" section. It includes a "File:" field with a "Choose File" button and the filename "reminder.html". The "Encoding:" dropdown is set to "iso-8859-1 (Western Europe)". The "Doctype:" dropdown is set to "HTML 4.01 Transitional". There are several checkboxes for options: "Use Fallback instead of Override" (checked), "Show Source", "Show Parse Tree", "Validate error pages", "Show Outline", "...exclude attributes", and "Verbose Output" (checked). A "Validate this file" button is at the bottom right. At the bottom of the form, there is a note: "If your document is on the Web, you can validate it with the same advanced options by [entering its address](#) instead." and a link to "Go to 'http://www.w3.org/'".

The HTML we have used is the sort you'll see on many web sites and in many applications; in fact, it's rather antiquated. At least it conforms to the transitional standard which supports the older tags and attributes that have been deprecated as well as the newer tags.

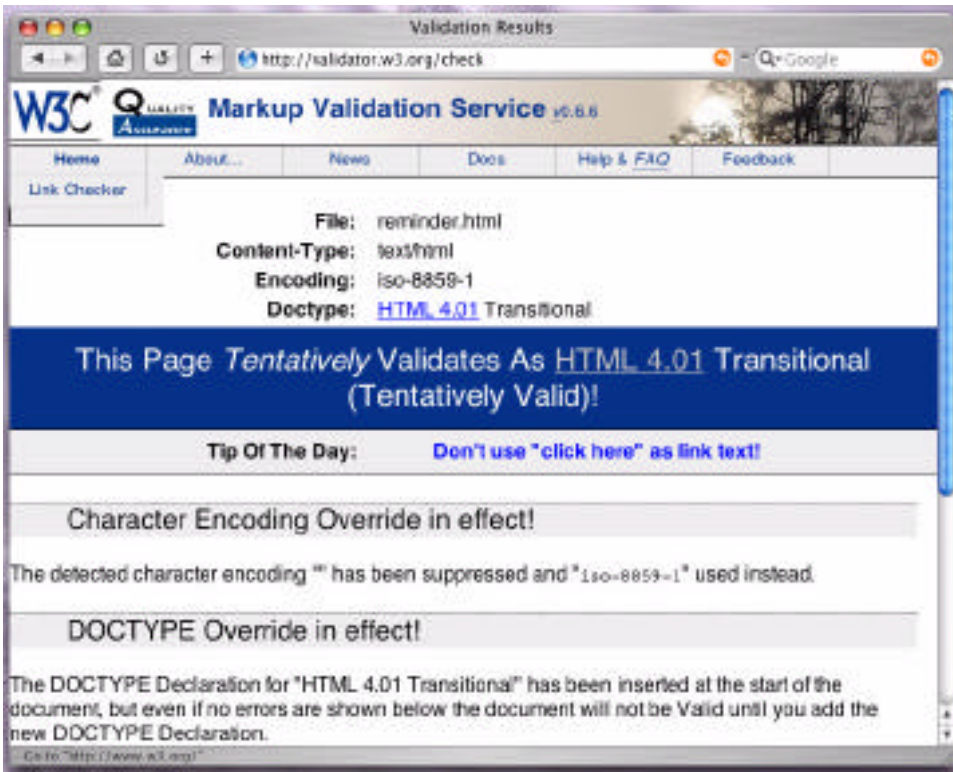


Figure 3 Our sample page, validated against the looser standard

However, if we are to use strict HTML 4.01 using only tags that have not been deprecated, some exceptions show up and our page is declared invalid:

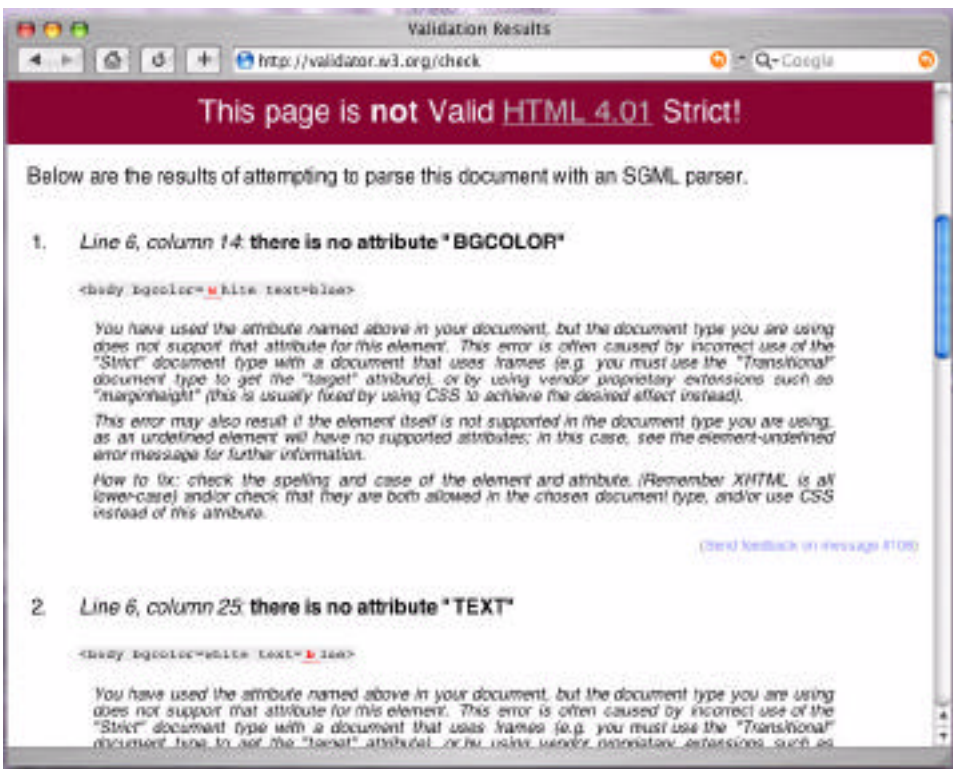


Figure 4 Our sample page, validated against the strict standard

# *License*

*These notes are distributed under the **Well House Consultants Open Training Notes License**. Basically, if you distribute it and use it for free, we'll let you have it for free. If you charge for its distribution of use, we'll charge.*

### 3.1 Open Training Notes License

Training notes distributed under the **Well House Consultants Open Training Notes License** (WHCOTNL) may be reproduced for any purpose PROVIDE THAT:

- This License statement is retained, unaltered (save for additions to the change log) and complete.
- No charge is made for the distribution, nor for the use or application thereof. This means that you can use them to run training sessions or as support material for those sessions, but you cannot then make a charge for those training sessions.
- Alterations to the content of the document are clearly marked as being such, and a log of amendments is added below this notice.
- These notes are provided "as is" with no warranty of fitness for purpose. Whilst every attempt has been made to ensure their accuracy, no liability can be accepted for any errors of the consequences thereof.

Copyright is retained by Well House Consultants Ltd, of 404, The Spa, Melksham, Wiltshire, UK, SN12 6QL - phone number +44 (1) 1225 708225. Email contact - Graham Ellis (graham@wellho.net).

Please send any amendments and corrections to these notes to the Copyright holder - under the spirit of the Open Distribution license, we will incorporate suitable changes into future releases for the use of the community.

If you are charged for this material, or for presentation of a course (Other than by Well House Consultants) using this material, please let us know. It is a violation of the license under which this notes are distributed for such a charge to be made, except by the Copyright Holder.

If you would like Well House Consultants to use this material to present a training course for your organisation, or if you wish to attend a public course is one is available, please contact us or see our web site - <http://www.wellho.net> - for further details.

Change log  
Original Version, Well House Consultants, 2004

Updated by: \_\_\_\_\_ on \_\_\_\_\_  
Updated by: \_\_\_\_\_ on \_\_\_\_\_  
Updated by: \_\_\_\_\_ on \_\_\_\_\_  
Updated by: \_\_\_\_\_ on \_\_\_\_\_  
Updated by: \_\_\_\_\_ on \_\_\_\_\_  
Updated by: \_\_\_\_\_ on \_\_\_\_\_

*License Ends.*