

## Teoría: Introducción al HTML

Antes de comenzar a hablar sobre el lenguaje con el que se crean las páginas Web (Html) vamos a comentar algunos conceptos generales sobre Internet.

### Índice de contenido

World Wide Web (WWW).....	1
¿Qué es HTML?.....	2
Versiones del HTML .....	3
Clientes y servidores WWW.....	4
Archivo HTML.....	4

## World Wide Web (WWW)

El servicio Web o WWW es una nueva forma de representar la información en Internet basada en páginas. Una página WWW puede incluir tres tipos de información: texto, gráficos e hipertexto.

Un *hipertexto* es texto resaltado que el usuario puede activar para cargar otra página WWW. La diferencia entre un documento hipertexto y un documento normal consiste en que el hipertexto contiene, además de la información, una serie de enlaces o conexiones con otros documentos relacionados, de manera que el lector puede pasar de un tema a otro y volver al documento original en el momento en que le interese.

Las **principales ventajas** del servicio WWW son tres.

- Se puede combinar texto y gráficos.
- Los hiperenlaces permiten cargar páginas de cualquier otro servidor conectado a Internet, da igual que esté localizado en España o en Australia.
- La creación de páginas WWW es bastante sencilla mediante el lenguaje HTML.

El gran éxito de Web no se debe solamente al empleo del hipertexto. Es normal encontrar que los documentos WWW están compuestos de texto y gráficos, y los enlaces con otros documentos pueden ser palabras clave subrayadas o resaltadas en el texto y gráficos, y los enlaces con otros documentos pueden ser palabras clave subrayadas o resaltadas en el texto, pero también la totalidad de una imagen o incluso partes de ella (como un mapa "sensible", que permite acceder a información sobre una ciudad haciendo un "click" del ratón sobre un determinado detalle del

plano). Y las últimas versiones de los programas navegadores (y la mayor rapidez de las telecomunicaciones) permiten integrar en un mismo documento texto, gráficos, sonidos o, incluso animaciones de video. Estos documentos compuestos son los que reciben el nombre de hipertexto. (PC WORLD Junio 1995)

## ¿Qué es HTML?

HTML, HyperText Markup Language, es un lenguaje simple utilizado para crear documentos de hipertexto para WWW. No es un lenguaje de descripción de página como Postscript; HTML no permite definir de forma estricta la apariencia de una página, aunque una utilización algo desviada hace que se utilice en ocasiones como un lenguaje de presentación. Además, la presentación de la página es muy dependiente del browser (o programa navegador) utilizado: el mismo documento no produce el mismo resultado en la pantalla si se visualiza con un browser en modo línea, Mosaic o Netscape, o sea, HTML se limita a describir la estructura y el contenido de un documento, y no el formato de la página y su apariencia.

Una de las claves del éxito de WWW, aparte de lo atractivo de su presentación es sin duda, **su organización y coherencia**. Todos los documentos WWW comparten un mismo aspecto y una única interfaz, lo que facilita enormemente su manejo por parte de cualquier persona. Esto es posible porque el lenguaje HTML, en que están escritos los documentos, no solo permite establecer hipervínculos entre diferentes documentos, sino que es un "lenguaje de descripción de página" independiente de la plataforma en que se utilice. Es decir un documento HTML contiene toda la información necesaria sobre su aspecto y su interacción con el usuario, y es luego el browser que utilizemos el responsable de asegurar que el documento tenga un aspecto coherente, independientemente del tipo de estación de trabajo desde donde estemos efectuando la consulta.

Su simplicidad es tal que no es necesario utilizar un editor particular. Su gran permisividad exige rigor y atención en la estructura de documentos con el fin de que éstos se visualicen correctamente al margen del contexto y el browser utilizado.

Por tanto, como hemos visto, HTML es un lenguaje muy sencillo que nos permite preparar documentos Web insertando en el texto de los mismos una serie de marcas (tags) que controlan los diferentes aspectos de la presentación y comportamiento de sus elementos.

Para escribir HTML lo único que se necesita es un editor de texto ASCII. Las marcas o tags que controlan el comportamiento del documento son fragmentos de texto encerrados entre los signos "mayor que" y "menor que" (<marca>). Existen diferentes tipos de marcas: algunas controlan simplemente la presentación del texto del documento; otras, la forma en que se incluirán

en él imágenes; otras, finalmente, los hiperenlaces con documentos o con diferentes partes del mismo documento. Existen una serie de programas, llamados Editores Web, que ayudan en la elaboración de documentos HTML, en nuestro caso utilizaremos el KOMPOZER (freeware), si bien, algunos elementos como los marcos los necesitaremos crear a mon con las etiquetas Web. Lo que si es necesario es un programa cliente WWW, tal como Explorer, o Firefox, para probar el documento a medida que lo vamos desarrollando.

Las marcas funcionan muchas veces por parejas, una para indicar el inicio de enlace o formato, y otra para señalar el final. La marca de inicio consiste en una letra o una palabra (por ejemplo, estas son marcas de inicio: <B>, <TITLE>). La marca de final es la misma letra o palabra precedida por la barra inclinada o "slash" (es decir, </B>, </TITLE>). Existen, no obstante, algunas marcas que no requieren su pareja de cierre, como <BR> (que fuerza un salto de línea). Es importante señalar que las marcas, en general pueden estar indistintamente en mayúsculas o en minúsculas.

## **Versiones del HTML**

Cuando se produjo la explosión de Internet el estándar de HTML que circulaba era el 2.0 (establecido en noviembre del 95), de modo que cualquier navegador que encontremos será capaz de interpretarlo. Prácticamente todo lo que veremos está contemplado por este estándar.

Aunque la versión 2.0 cumplía bien el objetivo para el que se creó, carecía de herramientas para tener un control mínimamente complejo de los documentos. Debido a ello, *Netscape* (líder del mercado de navegadores por aquel entonces) empezó a incluir etiquetas nuevas no incluidas en ningún estándar.

Por estos problemas, se creao un nuevo comité llamado W3C, que es el que actualmente elabora las nuevas versiones del HTML. Su primer trabajo consistió en crear el borrador del HTML 3.2, que incluía muchas de las mejoras que los principales navegadores (Netscape y Explorer) habían introducido en Internet, como son las tablas, los applets, etc..

Este borrador fue aprobado en enero de 1997 e inmediatamente el W3C se puso a trabajar en la elaboración del siguiente estándar: el 4.0.

En principio cualquier navegador debería ser capaz de interpretar el HTML 3.2.

## Cientes y servidores WWW

Para poder utilizar el servicio Web se necesitan dos partes. Por un lado, la empresa o institución que quiere facilitar su información tiene que crear páginas WWW, siguiendo el estándar definido por el lenguaje HTML, y ponerlas a disposición del público en Internet, en lo que se llama un servidor WWW. Por otro lado, el usuario que quiere acceder a dichas páginas tiene que utilizar un programa (cliente WWW) que lea las páginas WWW e interprete su significado (por ejemplo, un hipertexto). Estos programas *navegadores* o clientes WWW son los que permiten al ordenador del usuario interpretar el lenguaje HTML.

Los navegadores más usados son Explorer, *Firefox*, Chrome, etc.

## Archivo HTML

El lenguaje **HTML** se utiliza para definir un documento que se visualizará a través del navegador. Este documento se guarda en un archivo con extensión .htm ó .html

Básicamente los documentos escritos en HTML constan de texto y etiquetas. Las etiquetas permiten definir el formato del texto, el título que mostrará en la barra de título del navegador, los elementos multimedia que aparecerán incrustados en el documento pero que se almacenan en archivos externos, etc.

Todo documento HTML tiene la siguiente estructura: cabecera y cuerpo del documento. Las tres etiquetas que describen su estructura general son:

- `<html>`: indica que se inicia el documento.
- `<head>`: incluye el título de la página (`<title>`) que se muestra en la barra de título del navegador.

```
1. <!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.d
2. <html>
3. <head>
4.   <meta content="text/html; charset=ISO-8859-1"
5.     http-equiv="content-type">
6.   <title></title>
7. </head>
8. <body>
9.   <br>
10. </body>
11. </html>
```