

Fundamentos del HTML

Qué es el HTML

Estas siglas significan *Hypertext Markup Language*, que podría traducirse como “Lenguaje de Descripción de Hipertexto”. El HTML es una manera de describir documentos que contengan texto con formato, imágenes y vínculos a otros documentos. Es un subconjunto de un lenguaje más potente llamado SGML (*Standard Generalized Markup Language*).

Su especificación fue escrita por primera vez en 1991 por Tim Berners-Lee del CERN como parte de la iniciativa de creación del World Wide Web para facilitar la comunicación entre científicos (sobre todo físicos).

Estructura

Un documento HTML es un simple archivo de texto, que se almacena con extensión `html` o `htm`.

Al no ser más que un archivo de texto, se puede crear y modificar con cualquier programa editor de textos, pero para apreciar su aspecto final es necesario un programa visualizador: los navegadores de Internet.

En el archivo HTML se indican, mediante marcas específicas (las “etiquetas” o *tags*), el aspecto que debe tener la información textual, dónde deben aparecer las imágenes, etc. El programa visualizador utiliza estas marcas para realizar su labor.

Versiones

En pocos años de vida, el HTML ha evolucionado rápidamente; como es habitual, las versiones van numeradas. La diferencia entre versiones es que contemplan diferentes etiquetas; algunas quedan obsoletas de una versión a la siguiente, y las versiones superiores incorporan nuevas etiquetas con posibilidades más atractivas.

La última versión del HTML es la 4.01, cuya definición se hizo pública el 24 de diciembre de 1999. En este libro se tratará la versión 4.0, que fue revisada por segunda y última vez el 24 de abril de 1998.

Los navegadores están programados para interpretar una determinada versión de HTML, pero si se intenta visualizar un documento preparado según una versión superior, el visualizador debería ignorar elegantemente las marcas que no reconozca.

Limitaciones

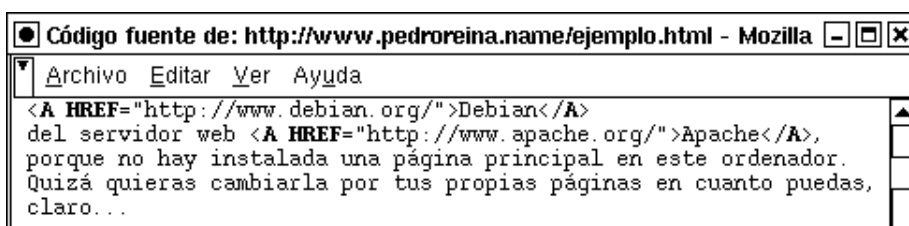
Existen una gran cantidad de parámetros que tienen importancia cuando se crean documentos pero que no se pueden definir en HTML (o resulta muy difícil hacerlo) por lo que no es un lenguaje muy apreciado por los tipógrafos. Su versatilidad y sencillez compensan esas limitaciones.

Distintos navegadores pueden mostrar el mismo archivo HTML de modos ligeramente diferentes, e incluso el mismo navegador puede ser configurado de distinta forma por cada usuario y así cambiar el aspecto del mismo archivo. Así pues, quién diseñe una página web debe asumir que podrá ser visualizado de muy diferentes formas.

Algunos visualizadores admiten etiquetas que sólo pueden interpretar ellos, con lo que se rompe la ideal universalidad de Internet. Es común indicar en la página cuál es el visualizador más indicado para ella.

Ver código de páginas Web

Una de las mejores maneras de aprender a usar el HTML es estudiar cómo ha sido usado en páginas que resulten de interés. En Mozilla Navigator, cuando se tiene visualizada una página Web basta elegir en el menú **Ver** la opción **Código fuente de la página** para que se abra una ventana que muestra el archivo tal como se escribió.



Entidad de referencia

Existen muchas extensiones diferentes de la especificación HTML, sobre todo de Netscape y Microsoft. Por eso es bueno que exista una organización que vele por la universalidad del lenguaje. Esa organización es el **World Wide Web Consortium**, cuya presentación se ve aquí:

