



ELEARNING TOTAL

Programador Web / Nivel 1 – Unidad 2

# Programador Web – Nivel 1

Unidad 2: Introducción a HTML





## *Indice*

### **Unidad 1: Introducción a HTML**

Características básicas

Etiquetas HTML



## Objetivos

### Que el alumno logre:

- Reconocer e implemente las diferentes etiquetas de HTML.



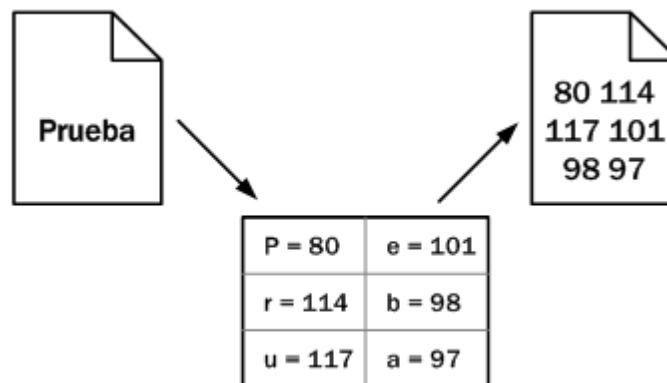


# Características básicas

## LENGUAJES DE ETIQUETAS

Uno de los retos iniciales a los que se tuvo que enfrentar la informática fue el de cómo almacenar la información en los archivos digitales. Como los primeros archivos sólo contenían texto sin formato, la solución utilizada era muy sencilla: se codificaban las letras del alfabeto y se transformaban en números.

De esta forma, para almacenar un contenido de texto en un archivo electrónico, se utiliza una tabla de conversión que transforma cada carácter en un número. Una vez almacenada la secuencia de números, el contenido del archivo se puede recuperar realizando el proceso inverso.



El proceso de transformación de caracteres en secuencias de números se denomina codificación de caracteres y cada una de las tablas que se han definido para realizar la transformación se conoce con el nombre de páginas de código. Una de las codificaciones más conocidas (y una de las primeras que se publicaron) es la codificación ASCII.

Una vez resuelto el problema de almacenar el texto simple, se presenta el reto de almacenar los contenidos de texto con formato. En otras palabras, ¿cómo se almacena un texto en negrita? ¿Y un texto de color rojo? ¿Y otro texto azul, en negrita y subrayado?



Utilizar una tabla de conversión similar a las que se utilizan para textos sin formato no es posible, ya que existen infinitos posibles estilos para aplicar al texto. Una solución técnicamente viable consiste en almacenar la información sobre el formato del texto en una zona especial reservada dentro del propio archivo. En esta zona se podría indicar dónde comienza y dónde termina cada formato.

No obstante, la solución que realmente se emplea para guardar la información con formato es mucho más sencilla: el archivo electrónico almacena

Tanto los contenidos como la información sobre el formato de esos contenidos. Si por ejemplo se quiere dividir el texto en párrafos y se desea dar especial importancia a algunas palabras, se podría indicar de la siguiente manera:

**<parrafo>**

**Contenido de texto con <importante>algunas palabras</ importante>  
resaltadas de forma especial.**

**</parrafo>**

El principio de un párrafo se indica mediante la palabra <párrafo> y el final de un párrafo se indica mediante la palabra </párrafo>. De la misma manera, para asignar más importancia a ciertas palabras del texto, se encierran entre <importante> y </importante>.

El proceso de indicar las diferentes partes que componen la información se denomina marcar (markup en inglés). Cada una de las palabras que se emplean para marcar el inicio y el final de una sección se denominan etiquetas.

Aunque existen algunas excepciones, en general las etiquetas se indican por pares y se forman de la siguiente manera:

**.: Etiqueta de apertura:** carácter <, seguido del nombre de la etiqueta (sin espacios en blanco) y terminado con el carácter >

**.: Etiqueta de cierre:** carácter <, seguido del carácter /, seguido del nombre de la etiqueta (sin espacios en blanco) y terminado con el carácter >



Así, la estructura típica de las etiquetas HTML es:

**`<nombre_etiqueta> ... </nombre_etiqueta>`**

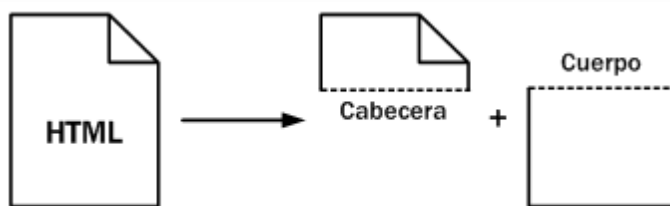
HTML es un lenguaje de etiquetas (también llamado lenguaje de marcado) y las páginas web habituales están formadas por cientos o miles de pares de etiquetas. De hecho, las letras “ML” de la sigla HTML significan “markup language”, que es como se denominan en inglés a los lenguajes de marcado. Además de HTML, existen muchos otros lenguajes de etiquetas como XML, SGML, DocBook y MathML.

La principal ventaja de los lenguajes de etiquetas es que son muy sencillos de leer y escribir por parte de las personas y de los sistemas electrónicos. La principal desventaja es que pueden aumentar mucho el tamaño del documento, por lo que en general se utilizan etiquetas con nombres muy cortos.



## EL PRIMER DOCUMENTO HTML

Las páginas HTML se dividen en dos partes: la **cabecera** y el **cuerpo**. La cabecera incluye información sobre la propia página, como por ejemplo su título y su idioma. El cuerpo de la página incluye todos sus contenidos, como párrafos de texto e imágenes.



El cuerpo (llamado body en inglés) contiene todo lo que el usuario ve en su pantalla y la cabecera (llamada head en inglés) contiene todo lo que no se ve (con la única excepción del título de la página, que los navegadores muestran como título de sus ventanas).

A continuación, se muestra el código HTML de una página web muy sencilla:

```
<!DOCTYPE html>
<html>
<head>
<title>El primer documento HTML</title>
</head>
<body>
<p>Contenido de nuestro primer document de HTML</p>
</body>
</html>
```

Si quieres probar este primer ejemplo, debes hacer lo siguiente:

1. Abrir un editor de archivos de texto y crea un archivo nuevo
2. Copiar el código HTML mostrado anteriormente y pégalo tal cual en el archivo que has creado
3. Guardar el archivo con el nombre que quieras, pero con la extensión .html



Para que el ejemplo anterior funcione correctamente, es imprescindible utilizar un editor de texto sin formato. Si tu sistema operativo es Windows, puedes utilizar el Bloc de notas, Wordpad, EmEditor, UltraEdit, Notepad++, Sublime Text3, Visual Code, etc. pero no puedes utilizar un procesador de textos como **Microsoft Word** u **Open Office**.

Si utilizas sistemas operativos tipo Linux, puedes utilizar editores como Gedit, Kedit, Kate e incluso Vi, pero no utilices KOffice ni Open Office.

Si ya estás viendo tu primera página HTML en el navegador, prueba a pulsar sobre el menú Ver > Código fuente y podrás ver el código HTML de la página que está cargada en el navegador. De hecho, **puedes ver el código HTML de cualquier página de Internet mediante la opción Ver > Código fuente**. Prueba a ver el código HTML de tu página preferida y verás cuantas etiquetas puede llegar a tener una página compleja.

Volviendo al código HTML del primer ejemplo, es importante conocer las cuatro etiquetas principales de un documento HTML (<html>, <head>, <body>):

**<!DOCTYPE html>:** La declaración doctype aparece justo encima de la etiqueta de <html>, al comienzo de cada documento. El doctype explica que tipo de HTML se espera, y por lo tanto, con cual validador de especificación (por ejemplo, [el validador de HTML del W3C](#)) debe validarse el documento. En este artículo del Currículo de estándares web vamos a explorar los diferentes doctypes con los que probablemente se encontrará en su viaje por la Web y analizaremos cómo difieren el XHTML y el HTML.

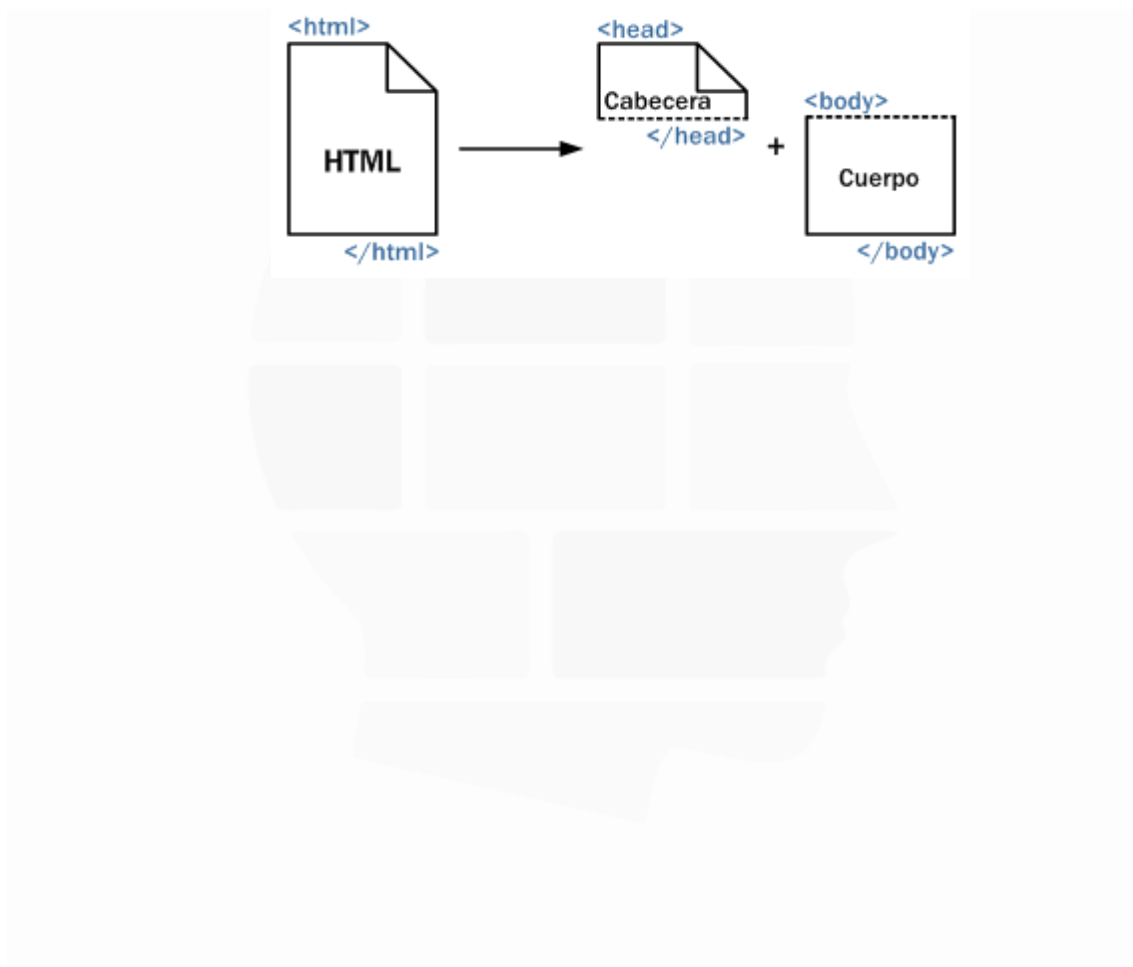
**<html></html>:** indica el comienzo y el final de un documento HTML. Ninguna etiqueta o contenido puede colocarse antes o después de la etiqueta <html> (con una sola excepción que se verá más adelante). En el interior de la etiqueta <html> se definen la cabecera y el cuerpo del documento HTML y todo lo que se coloque fuera de la etiqueta <html> se ignora.

**<head></head>:** delimita la parte de la cabecera del documento. La cabecera contiene información sobre el propio documento HTML, como por ejemplo su título y el idioma de la página. Los contenidos indicados en la cabecera no son visibles para el usuario, con la excepción de la etiqueta <title>, que se utiliza para indicar el título del documento y que los navegadores lo visualizan en la parte superior izquierda de la ventana del navegador (si no te has fijado anteriormente, vuelve a abrir el primer ejemplo en cualquier navegador y observa dónde se muestra el título de la página).





**<body></body>**: delimita el cuerpo del documento HTML. El cuerpo encierra todos los contenidos que se muestran al usuario (párrafos de texto, imágenes, tablas). En general, el <body> de un documento contiene cientos de etiquetas HTML, mientras que el <head> no contiene más que unas pocas.





## ETIQUETAS HTML

Todos los elementos del estándar HTML5 están listados en:

[https://developer.mozilla.org/es/docs/HTML/HTML5/HTML5\\_lista\\_elementos](https://developer.mozilla.org/es/docs/HTML/HTML5/HTML5_lista_elementos), descritos por su etiqueta de apertura y agrupados por su función.

El símbolo  indica que el elemento fue añadido en HTML5.

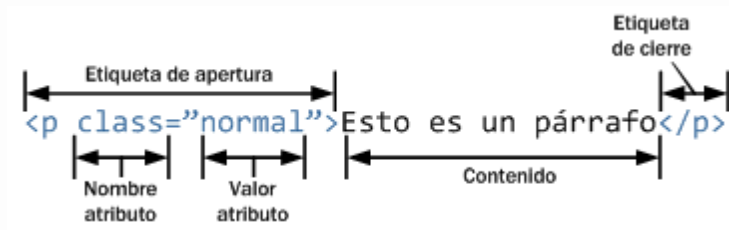
### Elementos HTML

Además de etiquetas y atributos, HTML define el término elemento para referirse a las partes que componen los documentos HTML.

Aunque en ocasiones se habla de forma indistinta de “elementos” y “etiquetas”, en realidad un elemento HTML es mucho más que una etiqueta, ya que está formado por:

- .: Una etiqueta de apertura.
- .: Cero o más atributos.
- .: Texto encerrado por la etiqueta.
- .: Una etiqueta de cierre.

El texto encerrado por la etiqueta es opcional, ya que algunas etiquetas de HTML no pueden encerrar ningún texto. El siguiente esquema muestra un elemento HTML, formado por una etiqueta `<p>`, atributos y contenidos de texto:



La estructura mostrada en el esquema anterior es un elemento HTML ya que comienza con una etiqueta de apertura (`<p>`), contiene cero o más atributos (`class="normal"`), dispone de un contenido de texto (Esto es un párrafo) y finaliza con una etiqueta de cierre (`</p>`). }



Por tanto, si una página web tiene dos párrafos de texto, la página contiene dos elementos y cuatro etiquetas (dos etiquetas <p> de apertura y dos etiquetas </p> de cierre). De todas formas, aunque estrictamente no son lo mismo, es habitual intercambiar las palabras “elemento” y “etiqueta”.

Por otra parte, el lenguaje HTML clasifica a todos los elementos en dos grupos: elementos en línea (inline elements en inglés) y elementos de bloque (block elements en inglés).

La principal diferencia entre los dos tipos de elementos es la forma en la que ocupan el espacio disponible en la página. Los elementos de bloque siempre empiezan en una nueva línea y ocupan todo el espacio disponible hasta el final de la línea, aunque sus contenidos no lleguen hasta el final de la línea. Por su parte, los elementos en línea sólo ocupan el espacio necesario para mostrar sus contenidos.

La mayoría de elementos de bloque pueden contener en su interior elementos en línea y otros elementos de bloque. Los elementos en línea sólo pueden contener texto u otros elementos en línea. En otras palabras, un elemento de bloque no puede aparecer dentro de un elemento en línea. En cambio, un elemento en línea puede aparecer dentro de un elemento de bloque y dentro de otro elemento en línea.

Los elementos en línea definidos por HTML son:

*a, abbr, acronym, b, basefont, bdo, big, br, cite, code, dfn, em, font, i, img, input, kbd, label, q, s, samp, select, small, span, strike, strong, sub, sup, textarea, tt, u, var.*

Los elementos de bloque definidos por HTML son:

*address, blockquote, center, dir, div, dlv, fieldset, form, h1, h2, h3, h4, h5, h6, hr, isindex, menu, noframes, noscript, ol, p, pre, table, ul.*

Los siguientes elementos también se considera que son de bloque:

*dd, dt, frame-set, li, tbody, td, tfoot, th, thead, tr.*

Los siguientes elementos pueden ser en línea y de bloque según las circunstancias:

*button, del, iframe, ins, map, object, script.*



## Resumen

### En esta Unidad...

En la presente unidad comenzamos a introducirnos en el lenguaje de creación de sitios web: HTML.

Con las etiquetas propuestas podemos comenzar a plantear la estructura semántica de una página web.

### En la próxima Unidad...

En la próxima unidad vamos a comenzar a trabajar con los nuevos elementos incorporados en la versión HTML5.