

## **Etiquetas y atributos**

HTML define varias etiquetas que los diseñadores pueden utilizar para marcar los diferentes elementos que componen una página, por ejemplo:

a, abbr, acronym, address, applet, area, b, base, basefont, bdo, big, block, quote, body, br, button, caption, center, cite, code, col, colgroup, dd, del, dfn, dir, div, dl, dt, em, fieldset, font, form, frame, frameset, h1, h2, h3, h4, h5, h6, head, hr, html, i, iframe, img, input, ins, isindex, kbd, label, legend, li, link, map, menu, meta, noframes, noscript, object, ol, optgroup, option, p, param, pre, q, s, samp, script, select, small, span, strike, strong, style, sub, sup, table, tbody, td, textarea, tfoot, th, thead, title, tr, tt, u, ul, var.

A pesar de que se trata de un número de etiquetas muy grande, no es suficiente para crear páginas complejas. Algunos elementos como las imágenes y los enlaces requieren cierta información adicional para estar completamente definidos. La etiqueta `<a>` por ejemplo se emplea para incluir un enlace en una página. Utilizando sólo la etiqueta `<a>` no es posible establecer la dirección a la que apunta cada enlace. Como no es viable crear una etiqueta por cada enlace diferente, la solución consiste en personalizar las etiquetas HTML mediante cierta información adicional llamada atributos. De esta forma, se utiliza la misma etiqueta `<a>` para todos los enlaces de la página y se utilizan los atributos para indicar la dirección a la que apunta cada enlace.

```
<!DOCTYPE html>
<html>
<head>
<title>Ejemplo de atributos en las etiquetas</title>
</head>
<body>
<p>
Los enlaces son muy fáciles de indicar:
<a>Soy un enlace incompleto, porque no tengo dirección de destino</a>.
<a href="http://www.google.com">Este otro enlace apunta a la página de Google</a>.
</p>
</body>
</html>
```

El primer enlace del ejemplo anterior no está completamente definido, ya que no apunta a ninguna dirección. El segundo enlace, utiliza la misma etiqueta `<a>`, pero añade información adicional mediante un atributo llamado href. Los atributos se incluyen dentro de la etiqueta de apertura. Por ahora no es importante comprender la etiqueta `<a>` ni el atributo href, ya que se explicarán con todo detalle más adelante. No todos los atributos se pueden utilizar en todas las etiquetas. Por ello, cada etiqueta define su propia lista de atributos disponibles. Además, cada atributo también indica el tipo de valor que se le puede asignar. Si el valor de un atributo no es válido, el navegador ignora ese atributo. Aunque cada una de las etiquetas HTML define sus propios atributos, algunos de los atributos son comunes a muchas o casi todas las etiquetas. De esta forma, es habitual explicar por separado los atributos comunes de las etiquetas para no tener que volver a hacerlo cada vez que se explica una nueva etiqueta.

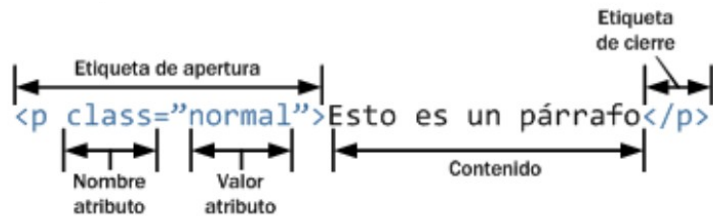
## **Elementos HTML**

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, dl, 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:

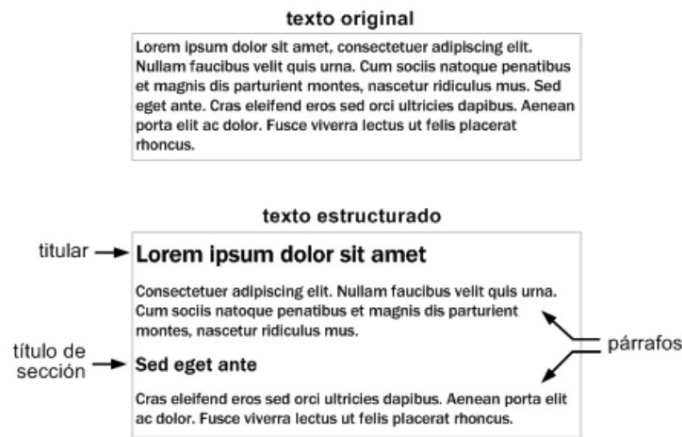
dd, dt, frame-set, li, tbody, td, tfoot, th, thead, tr.

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

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

## Texto

La mayor parte del contenido de las páginas HTML habituales está formado por texto, llegando a ser más del 90% del código de la página. Por este motivo, es muy importante conocer los elementos y etiquetas que define HTML para el manejo del texto. El lenguaje HTML incorpora al tratamiento del texto muchas de las ideas y normas establecidas en otros entornos de publicación de contenidos. De esta forma, HTML define etiquetas para estructurar el contenido en secciones y párrafos y define otras etiquetas para marcar elementos importantes dentro del texto. La tarea inicial del editor de contenidos HTML consiste en estructurar el texto original definiendo sus párrafos, titulares y títulos de sección, como se muestra en la siguiente imagen:



El anterior ejemplo muestra la transformación de un párrafo con un texto simple en un párrafo cuyo texto contiene elementos marcados de forma especial. Así, algunas palabras del texto se muestran en negrita porque se consideran importantes; otras palabras aparecen en cursiva, ya que se han marcado como destacadas e incluso una frase aparece tabulada y entre comillas, indicando que es una cita textual de otro contenido. En las secciones siguientes se muestran todas las etiquetas que define HTML para estructurar y marcar el texto. Además, se hace una mención especial al tratamiento que hace HTML de los espacios en blanco y las nuevas líneas. La forma más sencilla de estructurar un texto consiste en separarlo por párrafos. Además, HTML permite incluir títulos que delimitan cada una de las secciones.

## Párrafos

Una de las etiquetas más utilizadas de HTML es la etiqueta `<p>`, que permite definir los párrafos que forman el texto de una página. Para delimitar el texto de un párrafo, se encierra ese texto con la etiqueta `<p>`, como muestra el siguiente ejemplo:

```
<!DOCTYPE html>
<html>
<head>
<title>Ejemplo de texto estructurado con párrafos</title>
</head>
<body>
<p>Este es el texto que forma el primer párrafo de la página.
Los párrafos pueden ocupar varias líneas y el navegador se encarga de ajustar su longitud al
tamaño de la ventana.</p>
<p>El segundo párrafo de la página también se define encerrando su texto con la etiqueta p. El
navegador también se encarga de
separar automáticamente cada párrafo.</p>
</body>
</html>
```

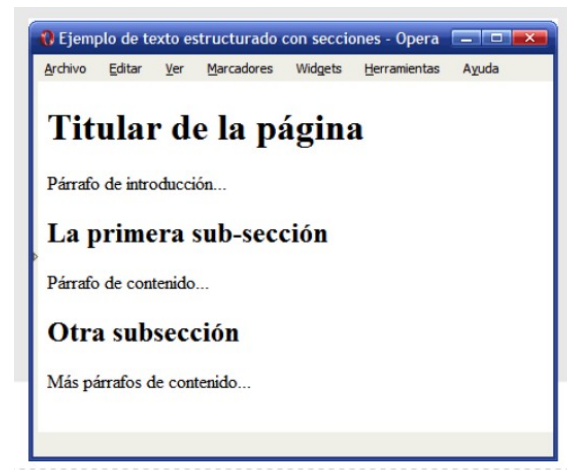
Los párrafos creados con HTML son elementos de bloque, por lo que siempre ocupan toda la anchura de la ventana del navegador. Además, no tienen atributos específicos, pero sí que se les pueden asignar los atributos comunes de HTML básicos, de internacionalización y de eventos.

## Encabezados

Las páginas HTML habituales suelen tener una estructura más compleja que la que se puede crear solamente mediante párrafos. De hecho, es habitual que las páginas se dividan en diferentes secciones jerárquicas o encabezados. Los títulos de sección se utilizan para delimitar el comienzo de cada sección de la página. HTML permite crear secciones de hasta seis niveles de importancia.

De esta forma, aunque una página puede definir cualquier número de secciones, sólo puede incluir seis niveles jerárquicos. Las etiquetas que definen los títulos de sección son <h1>, <h2>, <h3>, <h4>, <h5> y <h6>. La etiqueta <h1> es la de mayor importancia y por tanto se utiliza para definir los titulares de la página. La importancia del resto de etiquetas es descendiente, de forma que la etiqueta <h6> es la que se utiliza para delimitar las secciones menos importantes de la página. Las etiquetas <h1>, ..., <h6> definen títulos de sección, no secciones completas. Por este motivo, no es necesario encerrar los contenidos de una sección con su etiqueta correspondiente. Solamente se debe encerrar con las etiquetas <h1>, ..., <h6> los títulos de cada sección. El siguiente ejemplo muestra el uso de las etiquetas de título de sección:

```
<!DOCTYPE html>
<html>
<head>
<title>Ejemplo de texto estructurado con
secciones</title>
</head>
<body>
<h1>Titular de la página</h1>
<p>Párrafo de introducción...</p>
<h2>La primera sub-sección</h2>
<p>Párrafo de contenido...</p>
<h2>Otra subsección</h2>
<p>Más párrafos de contenido...</p>
</body>
</html>
```



Los navegadores asignan de forma automáticamente el tamaño del título de cada sección en función de su importancia. Así, los títulos de sección <h1> se muestran con el tamaño de letra más grande, ya que son el nivel jerárquico superior, mientras que los títulos de sección <h6> se visualizan con un tamaño de letra muy pequeño, adecuado para el nivel jerárquico de menor importancia. Evidentemente, el aspecto que los navegadores aplican por defecto a los títulos de sección se puede modificar utilizando las hojas de estilos de CSS.

### Marcado básico de texto

Una vez estructurado el texto en párrafos y secciones (encabezados), el siguiente paso es el marcado de los elementos que componen el texto. Los textos habituales están formados por elementos como palabras en negrita o cursiva, anotaciones y correcciones, citas a otros documentos externos, etc. HTML proporciona varias etiquetas para marcar cada uno de los diferentes tipos de texto. Entre las etiquetas más utilizadas para marcar texto se encuentran <em> y <strong>. La etiqueta <em> marca un texto indicando que su importancia es mayor que la del resto del texto. La etiqueta <strong> indica que un determinado texto es de la mayor importancia dentro de la página. Ejemplo:

```
<!DOCTYPE html>
<html>
<head>
<title>Ejemplo de etiqueta em y strong</title>
</head>
<body>
<p>El lenguaje HTML permite marcar algunos segmentos de texto como <em>muy
importantes</em> y otros segmentos como <strong>los más importantes</strong>.</p>
```

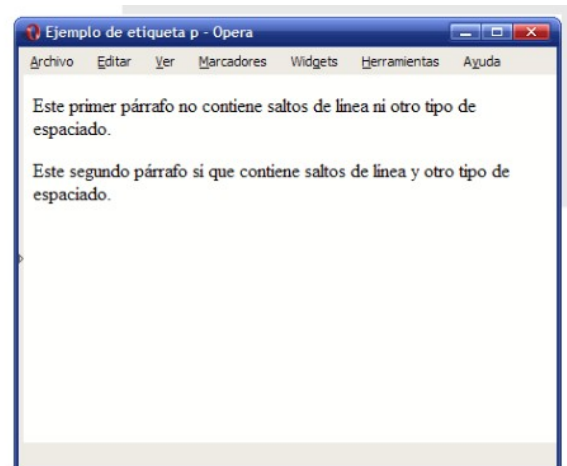
```
</body>
</html>
```

Por defecto, los navegadores muestran los elementos `<em>` en cursiva para hacer evidente su importancia y muestran los elementos `<strong>` en negrita, para indicar que son los más importantes.

## Espacios en blanco y nuevas líneas

El aspecto más sorprendente del lenguaje HTML cuando se desarrollan los primeros documentos es el tratamiento especial de los “espacios en blanco” del texto. HTML considera espacio en blanco a los espacios en blanco, los tabuladores, los retornos de carro y el carácter de nueva línea (ENTER o Intro).

```
<!DOCTYPE html>
<html>
<head><title>Ejemplo de etiqueta p</title></head>
<body>
<p>Este primer párrafo no contiene saltos de línea ni
otro tipo de espaciado.</p>
<p>Este segundo párrafo sí que contiene saltos
de
línea
y otro tipo de espaciado.</p>
</body>
</html>
```



Los dos párrafos de la imagen anterior se ven idénticos, aunque el segundo párrafo incluye varios espacios en blanco y está escrito en varias líneas diferentes. La razón de este comportamiento es que HTML ignora todos los espacios en blanco sobrantes, es decir, todos los espacios en blanco que no son el espacio en blanco que separa las palabras. No obstante, HTML proporciona varias alternativas para poder incluir tantos espacios en blanco y tantas nuevas líneas como sean necesarias dentro del contenido textual de las páginas.

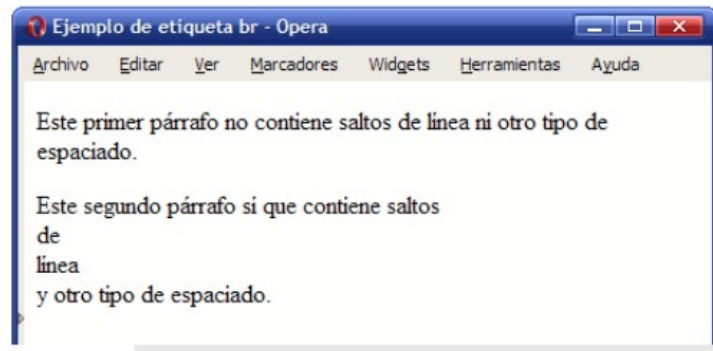
## Nuevas líneas

Para incluir una nueva línea en un punto y forzar a que el texto que sigue se muestre en la línea inferior, se utiliza la etiqueta `<br>`. En cierta manera, insertar la etiqueta `<br>` en un determinado punto del texto equivale a presionar la tecla ENTER (o Intro) en ese mismo punto. La etiqueta `<br>` es una de las pocas etiquetas especiales de HTML. La particularidad de `<br>` es que es una etiqueta vacía, es decir, no encierra ningún texto. De esta forma, la etiqueta debe abrirse y cerrarse de forma consecutiva: `<br></br>`. En estos casos, HTML permite utilizar un atajo para indicar que una etiqueta se está abriendo y cerrando de forma consecutiva: `<br/>` (también se puede escribir como `<br>`). Utilizando la etiqueta `<br>` se puede rehacer el ejemplo anterior para que respete las líneas que forman el segundo párrafo:

```
<html>
<head>
<title>Ejemplo de etiqueta br</title></head>
<body>
<p>Este primer párrafo no contiene saltos de línea ni otro tipo de espaciado.</p>
<p>Este segundo párrafo sí que contiene saltos <br>
```

```
de <br>
línea <br>
y otro tipo de espaciado.</p>
</body>
</html>
```

El navegador ahora sí que muestra correctamente las nuevas líneas que se querían insertar:



## Espacios en blanco

La solución al problema de los espacios en blanco no es tan sencilla como el de las nuevas líneas. Para incluir espacios en blanco adicionales, se debe sustituir cada nuevo espacio en blanco por el texto `&nbsp;` (es importante incluir el símbolo `&` al principio y el símbolo `;` al final). Así, el código HTML del ejemplo anterior se debe rehacer para incluir los espacios en blanco adicionales:

```
<!DOCTYPE html>
<html>
<head>
<title>Ejemplo de entidad &nbsp;</title>
</head>
<body>
<p>Este primer párrafo no contiene saltos de línea ni otro tipo de espaciado.</p>
<p>Este segundo párrafo sí que contiene saltos <br/>
de <br/>
línea <br/>
y &nbsp;&nbsp; otro &nbsp; tipo &nbsp; de &nbsp; espaciado.</p>
</body>
</html>
```

Cada texto `&nbsp;` solamente equivale a un espacio en blanco, por lo que se deben escribir tantos `&nbsp;` seguidos como espacios en blanco seguidos existan en el texto. Ahora el navegador sí que muestra correctamente los espacios en blanco (y las nuevas líneas) del segundo párrafo:

