

## TEMA 4. TEXTO

1. ESTRUCTURAR.....	3
1.1. PÁRRAFOS.....	3
1.2. ENCABEZADOS.....	4
2. MARCADO BÁSICO DE TEXTO.....	5
3. MARCADO AVANZADO DE TEXTO.....	9
4. ESPACIOS EN BLANCO Y NUEVAS LÍNEAS.....	10
4.1. NUEVAS LÍNEAS.....	11
4.2. ESPACIOS EN BLANCO.....	12
4.3. TEXTO PREFORMATEADO.....	13
5. CODIFICACIÓN DE CARACTERES.....	16

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.

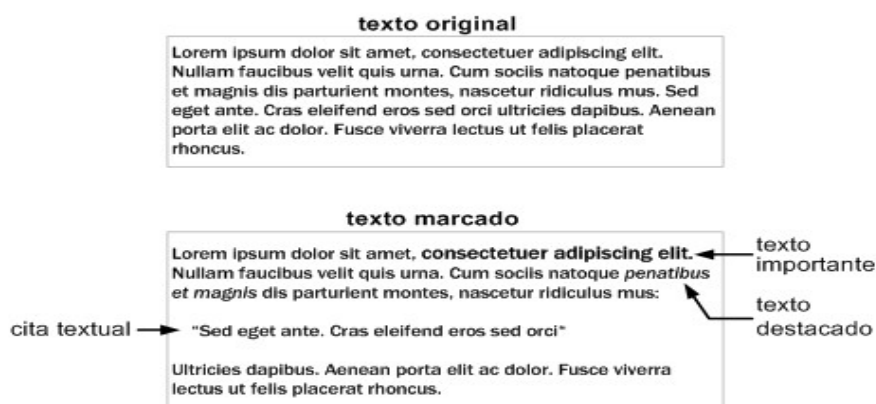
La tarea inicial del editor de contenidos HTML consiste en estructurar el texto original definiendo sus párrafos, titulares y títulos de sección, como se muestra en la siguiente imagen:



**Figura 4.1** Resultado de estructurar un texto sencillo

El proceso de estructurar un texto simple consiste en indicar las diferentes zonas o secciones que componen el texto. De esta forma, los textos estructurados utilizan etiquetas para delimitar cada párrafo y títulos de sección para delimitar cada una de las secciones que forman el texto.

Una vez definida la estructura básica de los contenidos de la página, el siguiente paso consiste en marcar los diferentes elementos dentro del propio texto: definiciones, abreviaturas, textos importantes, textos modificados, citas a otras referencias, etc.



**Figura 4.2** Resultado de marcar un texto sencillo

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.

## 1. ESTRUCTURAR

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.

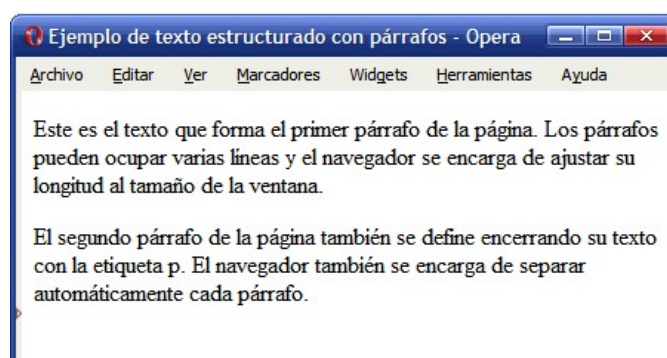
### 1.1. PÁRRAFOS

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:

```
<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>
```

El ejemplo anterior se visualiza de la siguiente manera en cualquier navegador:



**Figura 4.3** Ejemplo de texto HTML estructurado con párrafos

La siguiente tabla recoge la definición formal de la etiqueta `<p>`:

<b>Etiqueta</b>	<code>&lt;p&gt;</code>
<b>Tipo de elemento</b>	Bloque
<b>Descripción</b>	Delimita el contenido de un párrafo de texto

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.

## 1.2. 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.

Los encabezados se utilizan para delimitar el comienzo de una sección de la página. HTML permite crear encabezados 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 encabezados 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, normalmente sólo tendremos un encabezado `h1` por cada documento .html. 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.

A continuación, se muestra la definición formal de la etiqueta `<h1>`, siendo idéntica la definición del resto de etiquetas referidas a los títulos de sección:

Etiqueta	<code>&lt;h1&gt;</code>
Tipo de elemento	Bloque
Descripción	Define los títulos de las secciones de mayor importancia de la página.

Al igual que la etiqueta `<p>`, las etiquetas de título de sección son elementos de bloque y no tienen atributos específicos.

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:

```
<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 muestran el ejemplo anterior de la siguiente manera:

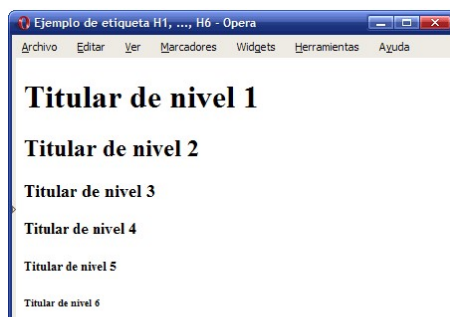


**Figura 4.4** Ejemplo de texto HTML estructurado con párrafos y secciones

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.

Es muy importante utilizar adecuadamente los encabezados, respetando el orden y la importancia que tiene cada uno de ellos, es decir, cuando utilicemos un encabezado `<h2>` será porque anteriormente hemos utilizado un encabezado `<h1>`. Es incorrecto utilizar un encabezado si anteriormente no hemos utilizado los encabezados superiores al que estamos utilizando.

Evidentemente, el aspecto que los navegadores aplican por defecto a los títulos de sección se puede modificar utilizando las hojas de estilos de CSS. La siguiente imagen muestra el tamaño por defecto con el que los navegadores muestran cada titular:



**Figura 4.5** Ejemplo de uso de las etiquetas h1, h2, h3, h4, h5 y h6

## 2. MARCADO BÁSICO DE TEXTO

Una vez estructurado el texto en párrafos y secciones, 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 definición formal de estas dos etiquetas se muestra a continuación:

<b>Etiqueta</b>	<code>&lt;em&gt;</code>
Tipo de elemento	En línea
Descripción	Realza la importancia del texto que encierra

Etiqueta	<code>&lt;strong&gt;</code>
Tipo de elemento	En línea
Descripción	Realza con la máxima importancia el texto que encierra

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:

```
<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:

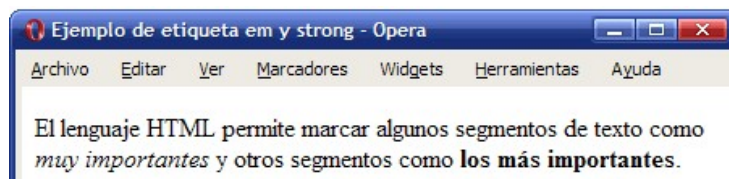
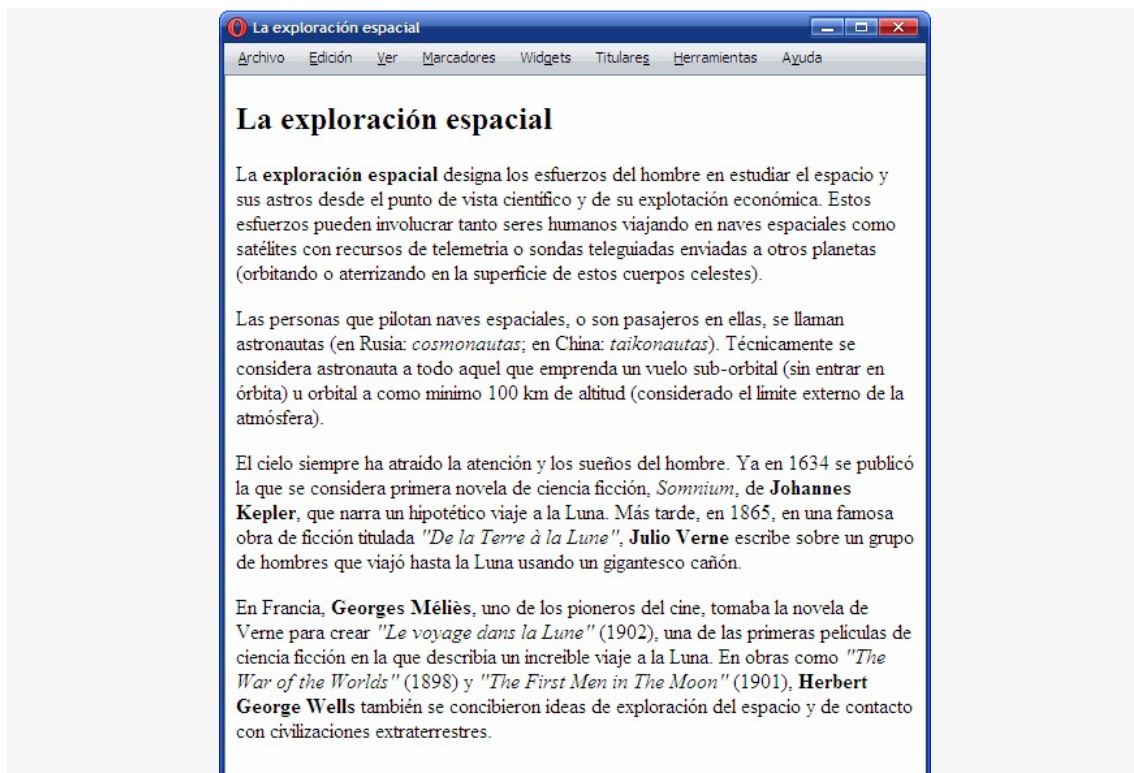


Figura 4.6 Ejemplo de uso de las etiquetas em y strong



**Figura 4.7** Resultado de estructurar y marcar el texto original

<b>Etiqueta</b>	<b>&lt;small&gt;</b>
<b>Tipo de elemento</b>	En línea
<b>Descripción</b>	Contenido irrelevante en el contexto de la página y únicamente informativo

```
<p><small>Acepto los términos de licencia</small></p>
```

HTML también permite marcar de forma adecuada las modificaciones realizadas en el contenido de una página. En otras palabras, HTML permite indicar de forma clara el texto que ha sido eliminado y el texto que ha sido añadido a un determinado texto original. Las etiquetas utilizadas son `<ins>` y `<del>`, cuya definición formal es la siguiente:

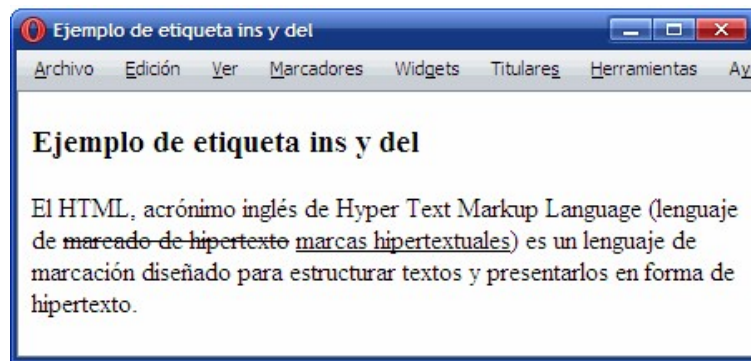
<b>Etiqueta</b>	<b>&lt;ins&gt;</b>
<b>Atributos propios</b>	<p><code>cite = "url"</code> - Indica la URL de la página en la que se puede obtener más información sobre el motivo por el que se realizó la modificación.</p> <p><code>datetime = "fecha"</code> - Especifica la fecha y hora en la que se realizó el cambio</p>
<b>Tipo de elemento</b>	Bloque y en línea
<b>Descripción</b>	Se emplea para marcar una modificación en los contenidos originales consistente en la inserción de un nuevo contenido
<b>Etiqueta</b>	<b>&lt;del&gt;</b>
<b>Atributos propios</b>	<p><code>cite = "url"</code> - Indica la URL de la página en la que se puede obtener más información sobre el motivo por el que se realizó la modificación.</p> <p><code>datetime = "fecha"</code> - Especifica la fecha y hora en la que se realizó el cambio</p>
<b>Tipo de elemento</b>	Bloque y en línea
<b>Descripción</b>	Se emplea para marcar una modificación en los contenidos originales consistente en el borrado de cierto contenido

Ejemplo:

```
<html>
  <head><title>Ejemplo de etiqueta ins y del</title></head>
  <body>
    <h3>Ejemplo de etiqueta ins y del</h3>
    <p>El HTML, acrónimo inglés de Hyper Text Markup Language
    (lenguaje de <del datetime="20091025"
    cite="http://librosweb.es/xhtml/capitulo_3/marcado_basico_de_te
    xto.html">marcado de hipertexto</del> <ins datetime="20091026"
    cite="http://librosweb.es/xhtml/capitulo_3.html"> marcas
    hipertextuales</ins>) es un lenguaje de marcación diseñado para
    estructurar textos y presentarlos en forma de hipertexto.</p>
  </body>
</html>
```

Los navegadores muestran el ejemplo anterior de la siguiente manera:





**Figura 4.8** Ejemplo de uso de las etiquetas ins y del

Por defecto, el texto eliminado (marcado con la etiqueta `<del>`) se muestra tachado de forma que el usuario pueda identificarlo fácilmente como un texto que formaba parte del texto original y que ya no tiene validez. El texto insertado (marcado con la etiqueta `<ins>`) se muestra subrayado, de forma que el usuario pueda identificarlo como un texto nuevo que no formaba parte del texto original.

Por otra parte, en muchos tipos de páginas (artículos, noticias) es habitual citar literalmente un texto externo. HTML define la etiqueta `<blockquote>` para incluir citas textuales en las páginas web. La definición de la etiqueta HTML con el nombre más largo se muestra a continuación:

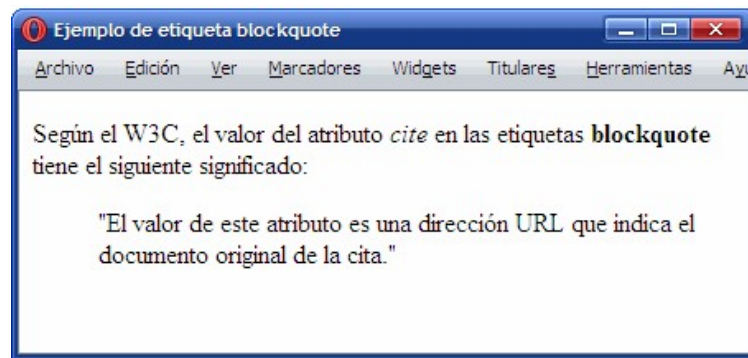
<b>Etiqueta</b>	<code>&lt;blockquote&gt;</code>
<b>Atributos propios</b>	<code>cite = "url"</code> - Indica la dirección de la página web original de la que se extrae la cita
<b>Tipo de elemento</b>	Bloque
<b>Descripción</b>	Se emplea para indicar que el texto que encierra es una cita textual de otro texto externo

Al igual que `<ins>` y `<del>`, la etiqueta `<blockquote>` permite indicar mediante el atributo `cite` la dirección de un documento del que se ha extraído la cita. Ejemplo:

```
<html>
  <head><title>Ejemplo de etiqueta blockquote</title></head>
  <body>
    <p>Según el W3C, el valor del
    atributo <em>cite</em> en las etiquetas
    <strong>blockquote</strong> tiene el
    siguiente significado:</p>
    <blockquote
      cite="http://www.w3.org/TR/html401/struct/text.html">El valor
      de este atributo
      es una dirección URL que indica el documento original de la
      cita.</blockquote>
  </body>
</html>
```

El aspecto que muestra el ejemplo anterior en cualquier navegador es el siguiente:





**Figura 4.9** Ejemplo de uso de la etiqueta blockquote

Para indicar de forma clara que el texto es una cita externa, los navegadores muestran por defecto el texto del elemento `<blockquote>` con un gran margen en la parte izquierda.

### 3. MARCADO AVANZADO DE TEXTO

Las páginas y documentos más avanzados suelen incluir otros elementos importantes que se deben marcar de forma adecuada. Por ello, HTML incluye muchas otras etiquetas que permiten marcar más elementos del texto.

La etiqueta `<abbr>` marca las abreviaturas de un texto y la etiqueta `<acronym>` se emplea para marcar las siglas o acrónimos del texto. Su definición es la siguiente:

<b>Etiqueta</b>	<code>&lt;abbr&gt;</code>
<b>Atributos propios</b>	<code>title = "texto"</code> - Indica el significado completo de la abreviatura
<b>Tipo de elemento</b>	En línea
<b>Descripción</b>	Se emplea para marcar las abreviaturas del texto y proporcionar el significado de esas abreviaturas

En ambos casos, el atributo `title` se puede utilizar para incluir el significado completo de la abreviatura o sigla. Ejemplo:

```
<html>
<head>
  <title>Ejemplo de etiqueta acronym</title>
</head>
<body>
  <p>
    Los <abbr title="Sistemas Informáticos">Sist. Inf. </abbr>.
  </p>
</body>
</html>
```

La mayoría de los navegadores muestran por defecto un borde inferior punteado para todos los elementos `<abbr>`. Al posicionar el puntero del ratón sobre la palabra subrayada, el navegador muestra un pequeño recuadro (llamado *tooltip* en inglés) con el valor del atributo `title`:

También podemos utilizarlos con iniciales.

Por otra parte, en ocasiones resulta útil incluir la definición de una palabra extraña o cuyo uso está restringido a un entorno muy determinado. HTML incluye la etiqueta `<dfn>` para proporcionar al usuario la definición de todas las palabras para las que se considere apropiado. La definición formal de esta etiqueta se muestra a continuación:

<b>Etiqueta</b>	<code>&lt;dfn&gt;</code>
<b>Atributos propios</b>	<code>title = "texto"</code> - Indica el significado completo del término
<b>Tipo de elemento</b>	En línea
<b>Descripción</b>	Se emplea para marcar las definiciones de ciertos términos y proporcionar el significado de esos términos

El siguiente ejemplo muestra cómo se utiliza la etiqueta `<dfn>` para incluir la definición completa de una palabra cuyo uso no es habitual fuera de los ámbitos médicos y psicológicos:

```
<p>Con estos síntomas, podría tratarse de un caso de
    <dfn title="Imagen o sensación subjetiva, propia de un sentido,
    determinada por otra sensación que afecta a un sentido
    diferente">sinestesia</dfn>
</p>
```

Por último, HTML incluye una etiqueta que se puede utilizar para marcar un texto como una citación:

<b>Etiqueta</b>	<code>&lt;cite&gt;</code>
<b>Tipo de elemento</b>	En línea
<b>Descripción</b>	Se emplea para marcar una cita o una referencia a otras fuentes

En ocasiones, no está clara la diferencia entre `<cite>` y `<blockquote>`. El elemento `<cite>` marca el autor de la cita (persona, documento, etc.) y `<blockquote>` marca el contenido de la propia cita. En el siguiente ejemplo, `<blockquote>` encierra el contenido de una frase célebre y `<cite>` encierra el nombre de su autor:

```
<p>Como dijo <cite>Mahatma Gandhi</cite>:
<blockquote>Vive como si fueras a morir mañana y aprende como si
fueras a vivir para siempre. </blockquote></p>
```

#### 4. 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).

El siguiente ejemplo ilustra este comportamiento:

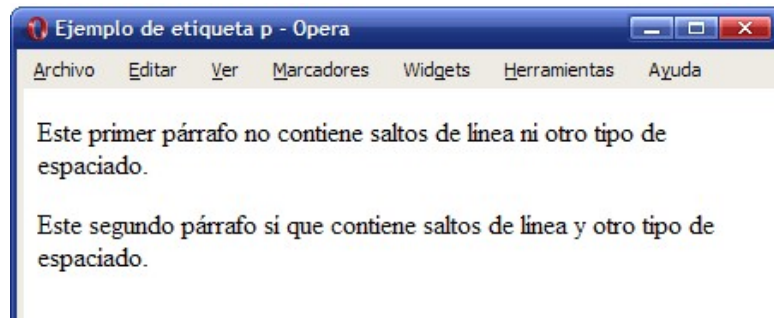
```
<html>
  <head><title>Ejemplo de etiqueta p</title></head>
  <body>
    <p>
      Este primer párrafo no contiene saltos de línea ni otro
      tipo de espaciado.
    </p>
    <p>Este segundo párrafo sí que contiene saltos
      de
      línea
```

```

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

```

El anterior código HTML se visualiza en cualquier navegador de la siguiente manera:



**Figura 4.12** Ejemplo de comportamiento de HTML con los espacios en blanco

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.

#### 4.1. 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 definición formal de `<br>` se muestra a continuación:

<b>Etiqueta</b>	<code>&lt;br&gt;</code>
<b>Tipo de elemento</b>	En línea y etiqueta vacía
<b>Descripción</b>	Fuerza al navegador a insertar una nueva línea

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>

```

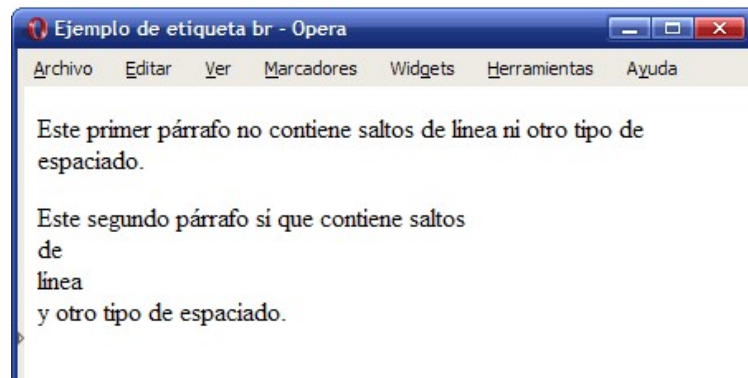
```

<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:



**Figura 4.13** Ejemplo de uso de la etiqueta br

#### 4.2. 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 `;` al final).

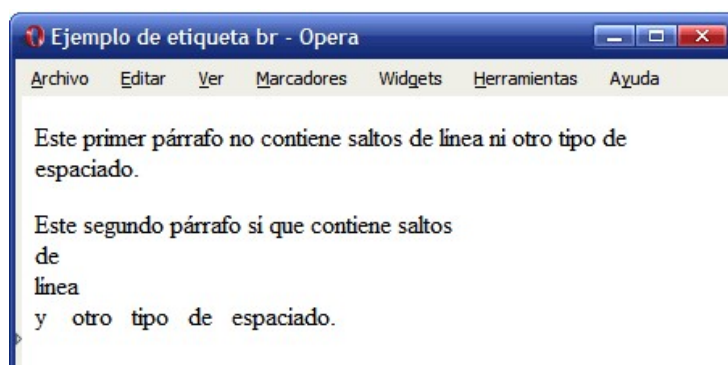
Así, el código HTML del ejemplo anterior se debe rehacer para incluir los espacios en blanco adicionales:

```

<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>

```

Ahora el navegador sí que muestra correctamente los espacios en blanco (y las nuevas líneas) del segundo párrafo:



**Figura 4.14** Ejemplo de uso de espacios en blanco en 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.

Más adelante se profundiza en el origen de `&nbsp;` y se comprenderá por qué es necesario incluir esa sucesión tan extraña de caracteres cada vez que se quiere incluir un espacio en blanco adicional.

Nombre	Diametro relativo	Periodo orbital	Número de lunas
Mercurio	0,382	0,24 años	0
Venus	0,949	0,62 años	0
Tierra	1	1 año	1
Marte	0,532	1,88 años	2
Júpiter	11,209	11,86 años	49
Saturno	9,449	29,46 años	52
Urano	4,007	84,01 años	27
Neptuno	3,883	164,80 años	13

**Figura 4.15** Texto HTML con espacios en blanco y nuevas líneas

### 4.3. TEXTO PREFORMATEADO

En ocasiones, es necesario mostrar los espacios en blanco de un texto que no se puede modificar. Se trata de un caso habitual cuando una página web debe mostrar directamente el texto generado por alguna aplicación.

En estos casos, se puede utilizar la etiqueta `<pre>`, que muestra el texto tal y como se ha escrito, respetando todos los espacios en blanco y todas las nuevas líneas. La definición formal de la etiqueta se muestra a continuación:

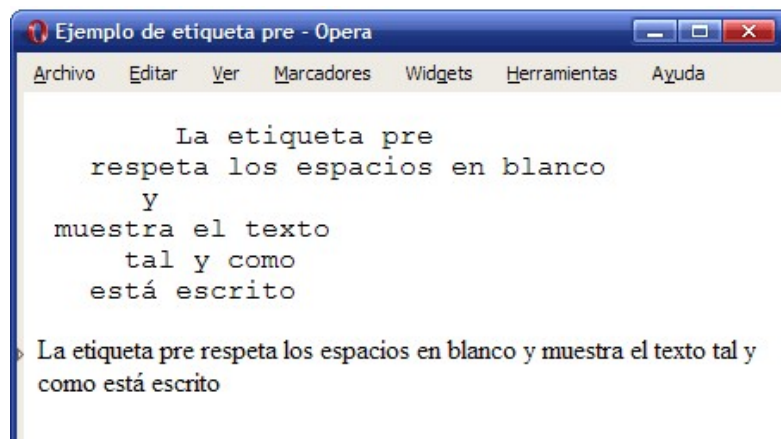
<b>Etiqueta</b>	<code>&lt;pre&gt;</code>
<b>Tipo de elemento</b>	Bloque
<b>Descripción</b>	Muestra el texto que encierra tal y como está escrito (respetando los espacios en blanco)

El siguiente ejemplo muestra el uso de la etiqueta `<pre>`:

```
<html>
```

```
<head><title>Ejemplo de etiqueta pre</title></head>
<body>
  <pre>
    La etiqueta pre
    respeta los espacios en blanco
    y
    muestra el texto
    tal y como
    está escrito
  </pre>
  <p>
    La etiqueta pre
    respeta los espacios en blanco
    y
    muestra el texto
    tal y como    está escrito
  </p>
</body>
</html>
```

El ejemplo anterior incluye el mismo texto (con espacios en blanco y varias líneas) dentro de una etiqueta `<pre>` y dentro de una etiqueta `<p>`. Las diferencias visuales en un navegador son muy evidentes:



**Figura 4.16** Ejemplo de uso de la etiqueta pre

El primer texto se ve en pantalla tal y como se ha escrito, respetando todos los espacios en blanco y todas las nuevas líneas. El segundo texto se ve como un párrafo normal, ya que HTML ha eliminado todos los espacios en blanco sobrantes.

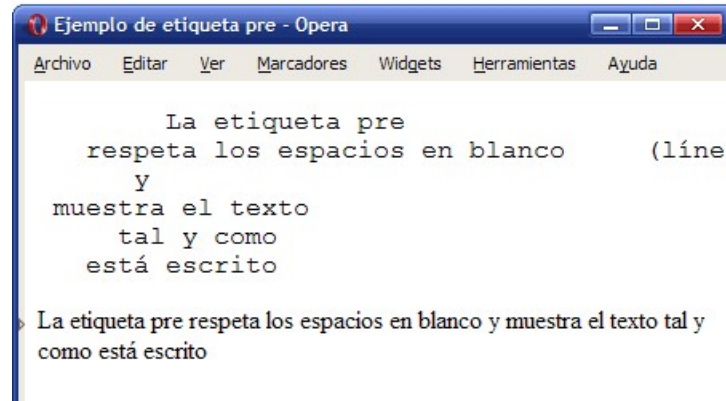
Los elementos `<pre>` son *especiales*, ya que los navegadores les aplican las siguientes reglas:

- Mantienen todos los espacios en blanco (tabuladores, espacios y nuevas líneas)
- Muestra el texto con un tipo de letra especial, denominado "*de ancho fijo*", ya que todas sus letras son de la misma anchura
- No se ajusta la longitud de las líneas (las líneas largas producen un *scroll* horizontal en la ventana del navegador)

Esta última característica diferencia por completo a los párrafos de los elementos `<pre>`. Como se ha visto, los navegadores ajustan la anchura de los párrafos de texto para que ocupen todo el tamaño de la ventana. Sin embargo, los elementos `<pre>` se muestran tal y como son originalmente, por lo que una línea muy larga dentro de un

elemento `<pre>` provoca que la anchura de la página sea superior a la anchura de la ventana del navegador.

Si en el ejemplo anterior se añade más texto al final de la segunda línea (para producir una línea larga), el navegador muestra un *scroll* horizontal ya que el texto completo no cabe en el tamaño de la ventana y las líneas de los elementos `<pre>` nunca se ajustan.



**Figura 4.17** Ejemplo de aparición de scroll horizontal con la etiqueta `pre`

Otra etiqueta relacionada con `<pre>` es la etiqueta `<code>`, que se utiliza para mostrar código fuente de cualquier lenguaje de programación. La definición formal de `<code>` es la siguiente:

<b>Etiqueta</b>	<code>&lt;code&gt;</code>
<b>Tipo de elemento</b>	En línea
<b>Descripción</b>	Delimita el texto considerado un fragmento de código fuente

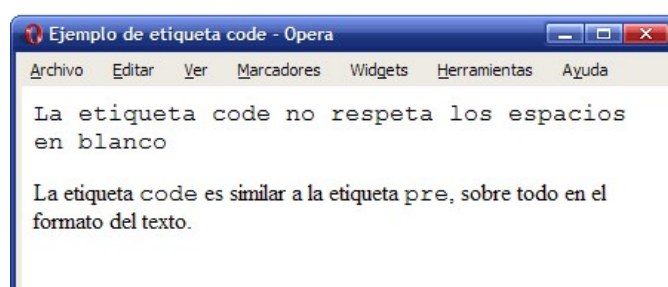
En la mayoría de las páginas web, no tiene sentido utilizar la etiqueta `<code>`. Sin embargo, en muchas páginas web técnicas que incluyen listados de programas en otros lenguajes, lo correcto es emplear la etiqueta `<code>`.

Ejemplo:

```
<html>
  <head><title>Ejemplo de etiqueta code</title></head>
  <body>
    <code>
      La etiqueta code
      no respeta los espacios en blanco
    </code>
    <p>La etiqueta <code>code</code> es similar a la
    etiqueta <code>pre</code>, sobre todo en el formato
    del texto.</p>
  </body>
</html>
```

El navegador muestra claramente el comportamiento de `<code>` y sus diferencias con `<pre>`:





**Figura 4.18** Ejemplo de uso de la etiqueta code

Al igual que sucede con los elementos `<pre>`, el texto encerrado por la etiqueta `<code>` se muestra con un tipo de letra especial de ancho fijo. Por el contrario, el elemento `<code>` no respeta los espacios en blanco ni las líneas, por lo que su comportamiento es similar a la etiqueta `<p>`. La última diferencia es que `<code>` es un elemento en línea, mientras que `<pre>` es un elemento de bloque.

## 5. CODIFICACIÓN DE CARACTERES

Una consideración importante directamente relacionada con el texto de las páginas HTML es la codificación de los caracteres y la inserción de caracteres *especiales*. Algunos de los caracteres que se utilizan habitualmente en los textos no se pueden incluir directamente en las páginas web:

- Los caracteres que utiliza HTML para definir sus etiquetas (`<`, `>` y `"`) no se pueden utilizar libremente.
- Los caracteres propios de los idiomas que no son el inglés (ñ, á, ç, è, i, etc.) pueden ser problemáticos dependiendo de la codificación de caracteres utilizada.

La solución a la primera limitación consiste en sustituir los caracteres reservados de HTML por unas expresiones llamadas *entidades HTML* y que representan a cada carácter:

Entidad	Carácter	Descripción	Traducción
<code>&amp;lt;</code>	<code>&lt;</code>	less than	signo de menor que
<code>&amp;gt;</code>	<code>&gt;</code>	more than	signo de mayor que
<code>&amp;amp;</code>	<code>&amp;</code>	ampersand	ampersand
<code>&amp;quot;</code>	<code>"</code>	quotation mark	comillas
<code>&amp;nbsp;</code>	(espacio en blanco)	non-breaking space	espacio en blanco
<code>&amp;apos;</code>	<code>'</code>	apostrophe	apóstrofo

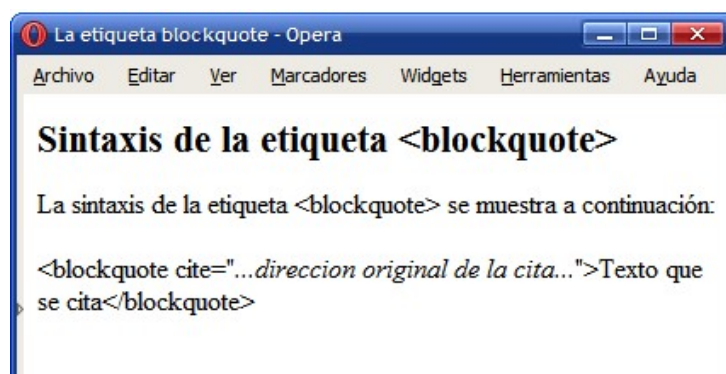
De esta forma, si se considera el siguiente texto:

Los caracteres `<`, `>`, `"` y `&` pueden dar problemas con los textos en HTML

Para mostrar correctamente el texto anterior en una página HTML, se debe sustituir cada carácter especial por su entidad HTML:

```
<p>Los caracteres &lt;; &gt;; &quot; y &amp; pueden dar problemas con los textos en HTML</p>
```

Determinar el código HTML que corresponde al siguiente documento:



**Figura 4.19** Texto HTML que incluye caracteres especiales

Por otra parte, los caracteres propios de los idiomas diferentes al inglés también pueden ser problemáticos. El motivo es que desde que se crea una página web hasta que llega al navegador del usuario, intervienen numerosos procesos:

- El diseñador crea la página web con su editor HTML (por ejemplo, Dreamweaver).
- Si se trata de una aplicación dinámica, el programador recorta la página HTML del diseñador y la mezcla con el resto del código de la aplicación (por ejemplo, PHP).
- El servidor web almacena las páginas HTML estáticas o el código de la aplicación web y sirve las páginas solicitadas por los usuarios.
- El usuario solicita y visualiza las páginas web a través de su navegador.

Si en todos los procesos anteriores se utiliza la misma codificación de caracteres, los caracteres propios de los idiomas se pueden escribir directamente:

```
<p>Este párrafo contiene caracteres acentuados y se almacena en formato UTF-8</p>
```

La palabra párrafo del ejemplo anterior incluye la letra á. Si el editor HTML del diseñador utiliza la codificación UTF-8, el entorno de desarrollo del programador también utiliza UTF-8, el servidor web sirve las páginas con esa codificación y el navegador del usuario es capaz de visualizar las páginas con formato UTF-8, el texto anterior se verá correctamente en el navegador del usuario.

Sin embargo, muchas veces no es posible que todos los procesos involucrados utilicen la misma codificación de caracteres. Por limitaciones técnicas o por decisiones de los diseñadores y programadores, los textos pueden pasar de codificación UTF-8 a codificación ISO-8859 en cualquier momento. Si se produce este cambio sin realizar una conversión correcta, el navegador del usuario mostrará caracteres extraños en todos los acentos y en todas las letras como la ñ.

La solución más sencilla para asegurar que todos estos caracteres potencialmente problemáticos se van a visualizar correctamente en el navegador del usuario consiste en sustituir cada carácter problemático por su entidad HTML:

Entidad	Carácter	Descripción oficial
&ntilde;	ñ	latin letter n with tilde
&Ntilde;	Ñ	latin capital n letter with tilde
&aacute;	á	a acute
&eacute;	é	e acute

&iacute;	í	i acute
&oacute;	ó	o acute
&uacute;	ú	u acute
&Aacute;	Á	A acute
&Eacute;	É	E acute
&Iacute;	Í	I acute
&Oacute;	Ó	O acute
&Uacute;	Ú	U acute
&euro;	€	euro

Así, el párrafo de texto del ejemplo anterior se podría escribir de la siguiente manera:

```
<p>Este párrafo contiene caracteres acentuados y se almacena en  
formato UTF-8</p>
```

Si se utilizan las entidades HTML en vez de los caracteres problemáticos, es indiferente pasar de una codificación de caracteres a otra diferente. En la Wikipedia se puede consultar la [lista completa de entidades HTML definidas](#).