

El objetivo de la presente lección es presentar cómo trabajar con texto en los documentos **HTML**, su razón principal de ser.

Comenzamos clasificando los elementos **HTML** en elementos de bloque o en línea. A continuación, presentamos el concepto de encabezamiento y qué elemento **HTML** utilizar para delimitarlo. Después, examinamos los principales elementos **HTML** para delimitar texto del contenido, tanto de bloque como en línea. Finalmente, mostramos qué es el pie y cómo delimitarlo. Y para terminar, las entidades **HTML**.

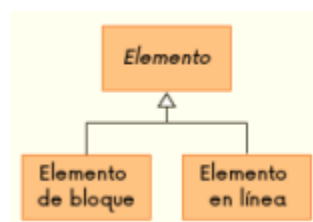
Al finalizar la lección, el estudiante sabrá:

- Qué es el encabezamiento y cómo delimitarlo mediante el elemento `<header>`.
- Qué es el pie y cómo delimitarlo mediante el elemento `<footer>`.
- Qué es un elemento de bloque.
- Qué es un elemento en línea.
- Qué elementos **HTML** utilizar para delimitar texto como, por ejemplo, `<p>`, `<em>`, `<strong>`, etc.
- Cuándo utilizar los elementos `<div>` y `<span>`.
- Cómo representar una línea horizontal.
- Cómo definir paneles.
- Cómo representar caracteres especiales mediante entidades.

## Introducción

Informalmente, se conoce como **elemento de texto** (*text element*) como aquel que hace referencia a texto del contenido como, por ejemplo, para dar vida a un párrafo, resaltar una determinada palabra o frase de un párrafo, etc.

Los elementos de **HTML** se pueden clasificar en elementos de bloque y en línea. Es importante tener claro que esta división es independiente de los elementos de texto. Es general al concepto de elemento.



Un **elemento de bloque** (*block element*) es aquel que representa una unidad, área o región de contenido. Y se presenta al usuario como un bloque, separado del anterior mediante un salto de línea. En cambio, un **elemento en línea** (*inline element*) es aquel que representa un fragmento de contenido dentro de un elemento, ya sea de bloque o bien de otro en línea. Este tipo de elementos se presentan dentro de un elemento de bloque.

Por ejemplo, el elemento `<p>`, con el que delimitar un párrafo, es un ejemplo de elemento en bloque. Mientras que el elemento `<em>`, con el que indicar un fragmento de texto enfatizado, es un elemento en línea:

```

<p>
  Esto es un ejemplo de párrafo.
  Dentro del cual, tenemos un <em>elemento en línea</em>.
</p>
  
```

Por convenio y buenas prácticas, generalmente un elemento de bloque tiene su etiqueta de inicio en

una línea; en otra, su contenido sangrado; y finalmente su etiqueta de fin, en otra aparte. Tal como muestra el ejemplo anterior.

## Encabezamientos

Una **encabezamiento** (*heading*) es un elemento de bloque con que se empieza otro elemento. Presenta contenido introductorio de una sección de contenido. Para delimitarlo, se recomienda utilizar un elemento `<header>`. Por favor, no confundirlo con `<head>`.

### Cabecera

El concepto de cabecera no es desconocido para nosotros. Cuando leemos un libro técnico o un artículo, por ejemplo, podemos ver distintas secciones mediante las cuales se encuentra estructurado el contenido. Todas ellas contienen cabeceras con las que se proporciona, principalmente, el título de la sección o del capítulo y/o una imagen. Los documentos **HTML** no son distintos, presentan su información en bloques, algunos de los cuales pueden presentar cabeceras.

Así pues, una **cabecera** (*header*) representa un encabezamiento, por ejemplo, de una sección, artículo o incluso del documento en sí, con contenido presentable al usuario, generalmente en forma de título o subtítulo. Las cabeceras se denotan mediante elementos de bloque `<header>`:

```
<header>Contenido</header>
```

Por lo general, las cabeceras tienen elementos de nivel de encabezado, párrafos, imágenes e incluso enlaces. He aquí un ejemplo:

```
<header>
  <h2>HTML5</h2>
  <h3>Especificación 5 del lenguaje HTML.</h3>
</header>
```

### Nivel de cabecera

**HTML** proporciona varios elementos con los que indicar el **nivel de cabecera** (*heading level*), esto es, la altura o profundidad del contenido dentro del documento. Al igual que los libros técnicos dividen cada capítulo en secciones y éstas, a su vez, en subsecciones y así sucesivamente, los documentos **HTML** también pueden desglosarse o dividirse en secciones, siendo el nivel de cabecera el que, por una parte, indica el título de la sección que marca y, por otra parte, denota en qué punto de la estructura del documento nos encontramos.

Los elementos de nivel de cabecera son `<h1>`, `<h2>`, `<h3>`, `<h4>`, `<h5>` y `<h6>`, donde el elemento `<h1>` es el que se encuentra en lo alto de la jerarquía o tabla de contenido. La sintaxis de estos elementos es:

```
<hx>Título</hx>
```

Ejemplo:

```
<h1>Elementos de texto</h1>
<p>...</p>

<h2>Introducción</h2>
<p>...</p>

<h2>Encabezamiento</h2>
<p>...</p>

<h3>Cabecera</h3>
<p>...</p>

<h3>Nivel de cabecera</h3>
<p>...</p>

<h2>Bloques de texto</h2>
<p>...</p>
```

Por razones **SEO**, todo documento debe tener una única cabecera `<h1>` y tantas `<h2>`-`<h6>` como sea necesario.

## Elementos de bloques de texto

Un **bloque de texto** (*text block*) es un área o unidad de texto en bloque. Se representa mediante elementos de bloque como `<p>`, `<pre>`, `<blockquote>`, etc.

### Párrafos

Un **párrafo** (*paragraph*) define una división o fragmento de contenido. Se representa mediante el elemento de bloque `<p>`:

`<p>Texto o contenido del párrafo</p>`

Dentro del párrafo, además de su texto o contenido, también puede aparecer otros elementos (en línea) como, por ejemplo, `<b>`, `<em>`, `<i>`, `<strong>`, `<span>`, etc. cuyo objeto es indicar una palabra o frase del párrafo con un significado especial.

He aquí un ejemplo ilustrativo:

```
<p>
Un <strong>sistema de gestión de bases de datos</strong> (<i>database management system</i>)
o <strong>SGBD</strong> (DBMS) es un software desarrollado para el almacenamiento,
organización y acceso de datos de manera sencilla y rápida. Existe cientos de sistemas de
gestión de bases de datos de diversa índole como, por ejemplo, ArangoDB, Cassandra, también
conocida como C*, MongoDB, Oracle, PostgreSQL, SQL Server o VoltDB. Concretamente,
<strong>PostgreSQL</strong> es un sistema de gestión de bases de datos objeto relacional, de
código abierto, desarrollado teniendo muy en cuenta el cumplimiento de estándares y su
extensibilidad. Se puede utilizar gratuitamente tanto para fines comerciales como privados.
Su página oficial es <a href="http://postgresql.org">postgresql.org</a>.
</p>
```

En los párrafos denotados mediante elementos `<p>`, las secuencias de dos o más espacios en blanco y los saltos de línea no se tienen en cuenta. Para el navegador, uno, dos o más espacios en blanco o saltos de línea se presentarán como un único espacio en blanco.

### Párrafo preformateado

Un **párrafo preformateado** (*preformatted paragraph*) es un bloque de texto cuyo contenido es literal y cuyos espacios en blanco y saltos de línea se interpretan como tales. Se representa mediante el elemento `<pre>`:

`<pre>Texto o contenido preformateado</pre>`

Al igual que el elemento `<p>` puede contener elementos en línea como, por ejemplo, `<em>`, `<strong>`, `<span>` o `<code>`.

### Cita

Una **cita** (*quote*) es un bloque o párrafo que contiene una cita o contenido que no es del propio autor, sino que procede de otro. Se representa mediante un elemento `<blockquote>`:

`<blockquote cite="url">Cita</blockquote>`

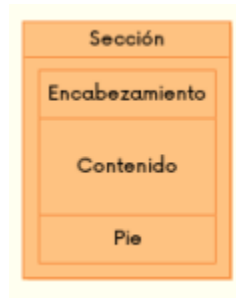
El atributo `cite` es opcional y, en caso de indicarse, señala dónde se encuentra la cita. Ejemplo:

```
<blockquote cite="http://www.w3.org/TR/1999/REC-html401-19991224/struct/text.html#edef-
BLOCKQUOTE">
  These two elements [Q and BLOCKQUOTE] designate quoted text.
</blockquote>
```

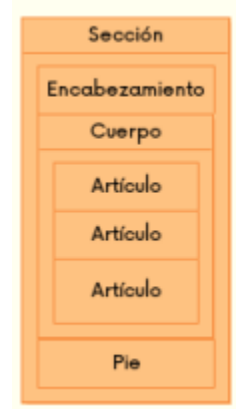
### Secciones

Una **sección** (*section*) representa una unidad o agrupación, generalmente temática, de contenido. Delimita una porción de contenido relacionado entre sí desde algún punto de vista. Por ejemplo, una sección de artículos, de comentarios, de *posts*, etc. Si es necesario, puede contener otras secciones.

La estructura principal de una sección consiste en un encabezamiento o cabecera, un cuerpo y un pie de contenido. El encabezamiento y el pie son opcionales, pudiendo haber secciones con sólo cuerpo, o sea, sólo de contenido.



El cuerpo o contenido de la sección generalmente está formado por otras secciones o artículos:



Ejemplo:

```
<section>
  <header>
    <h2>Comentarios</h2>
  </header>

  <article>Un comentario</article>
  <article>Otro comentario</article>
  <article>Otro comentario más</article>
</section>
```

## Artículos

Un **artículo** (*article*) representa una unidad independiente, distribuible y/o reusable de contenido como, por ejemplo, un *post*, una noticia, un comentario, una descripción, un producto, la información de un empleado, etc. Los artículos se representan mediante elementos `<article>`:

```
<article>Contenido</article>
```

Al igual que las secciones, un artículo puede tener encabezamiento, cuerpo y pie. Generalmente el encabezamiento se utiliza para proporcionar el título, la fecha de publicación, el autor, el nombre del producto, su fabricante, su vendedor, etc.

## División

Una **división** (*div*) representa un bloque de contenido que no puede declararse mediante ningún otro elemento de bloque definido por el estándar **HTML**. Se representa mediante un elemento `<div>`:

```
<div>Contenido</div>
```

Siempre que sea posible, hay que intentar utilizar el elemento **HTML** que mejor describa o represente un bloque de información. Dejando los elementos `<div>` para cualquier otro tipo no existente en el estándar **HTML**.

## Elementos en línea

Un **elemento en línea** (*inline element*) es aquel que delimita un fragmento de contenido dentro de otro, ya sea de bloque o también en línea. Se utilizan principalmente para marcar palabras o frases que deben presentarse de una manera especial, por ejemplo, en cursiva, negrita, coloreado, etc.

## Énfasis

Se habla de **énfasis** (*emphasis*) para representar un fragmento de texto sobre el que hay que prestar especial atención. Generalmente, los navegadores lo presentan en cursiva, pero no tiene por qué ser siempre así. Para marcar este tipo de texto, se utiliza el elemento en línea `<em>`:

```
<em>texto</em>
```

Ejemplo ilustrativo:

```
<p>A veces, todo va bien; otras veces no, pero <em>todo</em> tiene solución.</p>
```

Con un objeto similar, se puede utilizar el elemento en línea `<i>`, aunque éste se destina más para indicar una palabra técnica o en otro idioma:

```
<i>texto</i>
```

## Importancia

También se puede marcar un fragmento de texto para que se le preste especial atención, más relacionado con un nivel alto de importancia que con énfasis. Se utiliza el elemento en línea `<strong>` y se suele presentar en negrita, pero no siempre tiene por qué ser así:

```
<strong>texto</strong>
```

He aquí un ejemplo:

```
<p>
  Un <strong>artículo</strong> (<i>article</i>) representa una unidad independiente,
  distribuable y/o reusable de contenido como, por ejemplo, un post, una noticia, un
  comentario, una descripción, etc.
</p>
```

Si lo que se desea es resaltar algo, sin necesariamente darle importancia, se puede utilizar el elemento en línea `<b>`:

```
<b>texto</b>
```

## Texto pequeño

El elemento en línea `<small>` se utiliza para delimitar texto con un tamaño más pequeño que el estándar utilizado en el contenido principal de la página, por ejemplo, la leyenda de derechos de autor:

```
<small>texto</small>
```

Ejemplo:

```
<p>
  <small>Copyright &copy; 2016 nodemy.com. Reservados todos los derechos.</small>
</p>
```

## Código

Un elemento de **código** (*code*) es aquel que representa un fragmento de código como, por ejemplo, una expresión de un lenguaje de programación, el nombre de una función, método, clase, etc. Los navegadores suelen presentarlos en fuentes monoespacio. Se denota mediante el elemento en línea `<code>`:

```
<code>texto</code>
```

Ejemplo:

```
<p>Para marcar este tipo de texto, se utiliza el elemento en línea <code>em</code>.</p>
```

De manera similar, si se desea delimitar el nombre de una variable o parámetro se puede utilizar el elemento en línea especializado `<var>`:

```
<var>nombre variable</var>
```

También se puede utilizar el elemento en línea `<samp>` que delimita un ejemplo de salida de un programa:

```
<samp>salida</samp>
```

## Subíndices y superíndices

Un **subíndice** (*subscript*) es una letra o número que se coloca en la parte inferior derecha de una palabra a modo de índice que la distingue de otras. Mientras que un **superíndice** (*superscript*) se coloca

en la parte superior derecha. Los subíndices se delimitan mediante el elemento `<sub>` y los superíndices mediante `<sup>`:

```
<sub>índice</sub>
<sup>índice</sup>
```

He aquí un ejemplo ilustrativo de uso de superíndices:

$E = mc^2$

Lo anterior se presentaría al usuario como:

$E = mc^2$

## Abreviaturas y acrónimos

Una **abreviatura** (*abbreviation*) denota una versión reducida de una palabra mediante la supresión de sus letras finales. En cambio, un **acrónimo** (*acronym*) es un vocablo que denota una serie de palabras mediante la unión de los elementos iniciales de cada una de ellas. En **HTML**, se pueden representar ambos tipos de palabras mediante el elemento en línea `<abbr>`:

```
<abbr title="palabras">abreviatura o acrónimo</abbr>
```

Ejemplo:

```
<p>
  El sitio web oficial del <abbr title="World Wide Web Consortium">W3C</abbr> es
  w3.org.
</p>
```

## Fechas y horas

Las fechas y las horas son muy frecuentes en los sitios webs. Se utilizan para indicar el día y la hora de inicio o fin de un evento, la fecha de publicación de una noticia o de un *post*, la fecha desde la que se está vendiendo un producto, etc. En unos países, las fechas se representan de una manera y en otros, de otra; es más, cada autor puede utilizar un formato u otro. Como los motores de búsqueda no pueden comprender las manías de cada uno de nosotros, en la especificación **HTML5** se ha definido el elemento en línea `<time>`, con el que indicar dos cosas: por un lado, la fecha en un formato estándar y que será utilizada principalmente por los motores de búsqueda y, por otro lado, la fecha en el formato con el que presentársela a los usuarios.

La sintaxis del elemento `<time>` es como sigue:

```
<time datetime="fecha y/u hora en formato estándar">fecha y/u hora en formato usuario</time>
```

Los formatos de fecha y hora estándares más utilizados son los siguientes:

```
año-mes
año-mes-día
hora:minuto
hora:minuto:segundo
año-mes-día hora:minuto
año-mes-día hora:minuto:segundo
```

Ejemplo:

```
<time datetime="2016-09-27 08:31">27 sep'16 8h31m</time>
```

## span

El elemento en línea `<span>` es similar al elemento de bloque `<div>`, pero dentro de un bloque. Se utiliza para marcar un texto que no puede clasificarse mediante ninguno de los elementos en línea proporcionados por el estándar. Su sintaxis es:

```
<span>texto</span>
```

## Salto de línea

En **HTML**, la mayoría de los elementos no tienen en cuenta las secuencias de espacios en blanco, los tabulados y los saltos en línea. Cuando en un párrafo se indica varios espacios en blanco consecutivos, el navegador mostrará sólo uno. Si contiene un salto de línea, no lo muestra, lo trata como un espacio más. Si es necesario que el párrafo haga un salto de línea en un determinado punto del texto, se utiliza el elemento en línea `<br>`, cuya sintaxis es muy sencilla:

```
<br>
```

Por ejemplo, el siguiente párrafo:

```
<p>
  uno
  dos<br>
  tres<br>cuatro
</p>
```

Se mostrará como sigue:

```
uno dos
tres
cuatro
```

## Líneas horizontales

Una **línea horizontal** (*horizontal/ row*) es simplemente eso, una línea horizontal. Se suele utilizar para indicar un cambio de temática tras un párrafo. Se representa mediante el elemento `<hr>`:

```
<hr>
```

## Paneles

Un **panel** (*pane*) es un compartimento que contiene generalmente texto relacionado con el contenido principal de la página web, pero sin formar parte de él como, por ejemplo, un panel de noticias, un explorador del sitio web, un *banner*, etc. Generalmente, se representan mediante el elemento de bloque `<aside>`:

```
<aside>contenido</aside>
```

El elemento `<aside>` no se utiliza sólo para representar paneles, también se utiliza para texto separado del contenido principal como, por ejemplo, un panel incrustado dentro del cuerpo del documento que contenga una cita.

## Pies

Un **pie** (*footer*) es la contraparte del encabezado, cierra otro elemento, por ejemplo, de una sección, artículo o incluso del documento en sí. Se suele utilizar para indicar la información de contacto o derechos de autor. Se denota mediante el elemento de bloque `<footer>`:

```
<footer>Contenido del pie</footer>
```

Puede contener muchos tipos de elementos de bloque como, por ejemplo, `<nav>`, `<p>`, `<img>`, etc. y también en línea como `<b>`, `<i>`, etc. He aquí un ejemplo ilustrativo:

```
<footer>
  <p>Copyright &copy; 2016 <em>mi.sitio.web</em>. Reservados todos los derechos.</p>
</footer>
```

## Entidades

Algunos caracteres tienen un significado especial en **HTML**. Por ejemplo, el carácter `<` se utiliza para marcar el inicio de una etiqueta. ¿Cómo podemos añadir estos caracteres especiales al documento como texto puro? Por ejemplo, si deseamos mostrar el contenido *Con `<p>` se delimita un párrafo*, ¿cómo lo hacemos?

Está claro que lo siguiente no funcionará:

```
<p>Con <p> se delimita un párrafo</p>
```

Esto es lo que mostrará el navegador:

Con  
se delimita un párrafo

No hay señal ninguna de `<p>`. Lo trata como el elemento `<p>`, no como el texto `<p>` que deseamos mostrar al usuario.

Para resolver el problema, hay que utilizar lo que se conoce como **entidad** (*entity*), secuencia de caracteres que representa un determinado carácter. Estas secuencias comienzan por un *ampersand* (`&`) y terminan en punto y coma (`;`). Así por ejemplo, la entidad que representa el carácter `<` es `&lt;`. Por lo

tanto, nuestro ejemplo anterior debería escribirse como sigue:

`<p>Con &lt;p> se delimita un párrafo</p>`

Con `<p>` se delimita un párrafo

A continuación, se muestra las entidades más utilizadas:

| Entidad                 | Carácter |
|-------------------------|----------|
| <code>&amp;lt;</code>   | <        |
| <code>&amp;gt;</code>   | >        |
| <code>&amp;amp;</code>  | &        |
| <code>&amp;copy;</code> | ©        |