



Curso de XHTML

Unidad Didáctica 04: Texto



Ayuntamiento
de Vitoria-Gasteiz
Vitoria-Gasteizko
Udala

Índice de contenidos

- Introducción
- Cómo estructurar un texto
- Marcado avanzado de un texto
- Espacios en blanco y nuevas líneas
- Codificación de caracteres
- Conclusiones



Introducción

La mayor parte de las páginas HTML están formadas por texto (llegando a ser más del 90% del código de la página). Este lenguaje define tags para estructurar el contenido y otros para marcar elementos importantes dentro del texto.



Introducción

El proceso de estructuración de un texto simple consiste en indicar las diferentes zonas o secciones que componen un texto: como son los párrafos o títulos de sección.

El proceso posterior a estructurar el texto consiste en marcar los diferentes elementos dentro de éste: definiciones, abreviaturas, textos importantes, textos modificados, citas a otras referencias, etc.

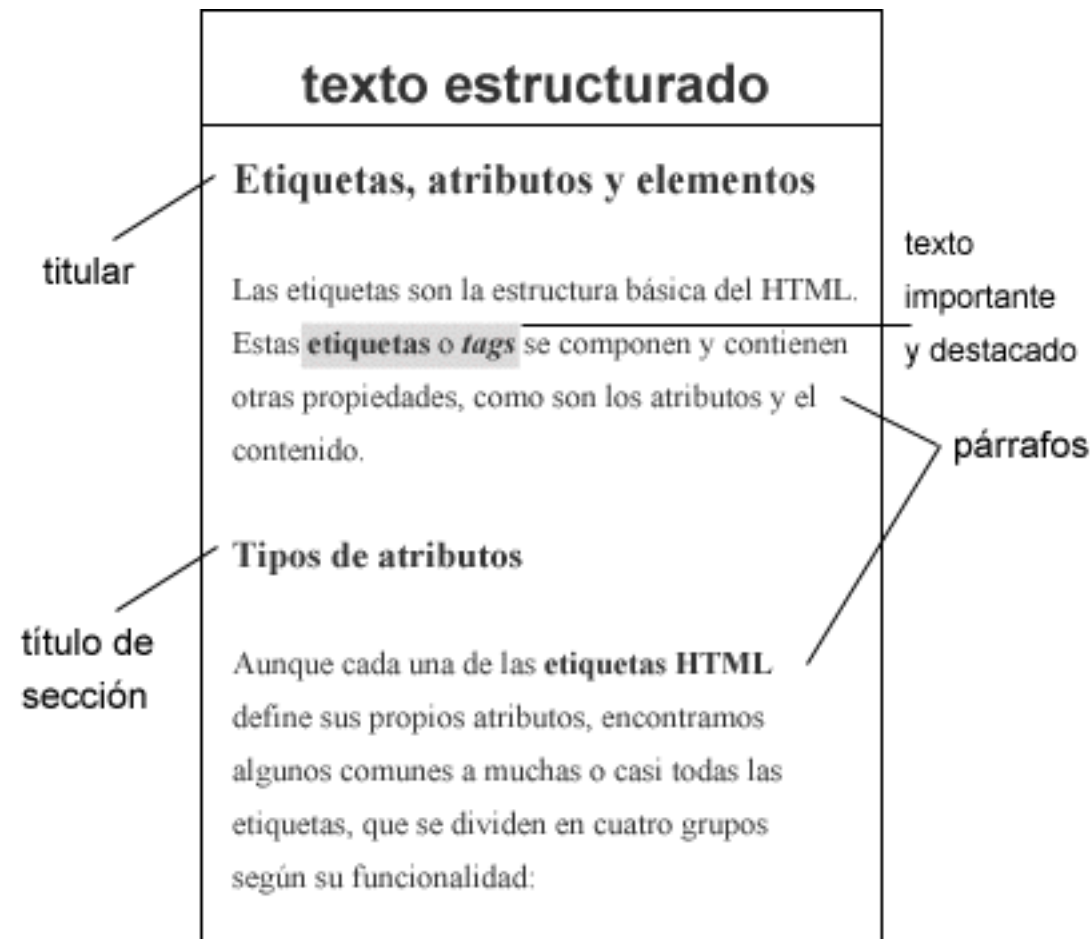


Introducción

texto original	texto estructurado
<p>Etiquetas, atributos y elementos. Las etiquetas son la estructura básica del HTML. Estas etiquetas o tags se componen y contienen otras propiedades, como son los atributos y el contenido. Tipos de atributos. Aunque cada una de las etiquetas HTML define sus propios atributos, encontramos algunos comunes a muchas o casi todas las etiquetas, que se dividen en cuatro grupos según su funcionalidad:</p>	<p>Etiquetas, atributos y elementos</p> <p>Las etiquetas son la estructura básica del HTML. Estas etiquetas o tags se componen y contienen otras propiedades, como son los atributos y el contenido.</p> <p>Tipos de atributos</p> <p>Aunque cada una de las etiquetas HTML define sus propios atributos, encontramos algunos comunes a muchas o casi todas las etiquetas, que se dividen en cuatro grupos según su funcionalidad:</p>



Introducción



Cómo estructurar un texto

PÁRRAFOS

Etiqueta <p>

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: -

Tipo de elemento: En bloque

Descripción: Delimita el contenido de un párrafo de texto



Cómo estructurar un texto

Esta etiqueta `<p>` permite definir los párrafos que forman el texto de una página. Como se puede ver en la tabla anterior, estos párrafos son elementos en bloque, por lo que ocupan toda la anchura del navegador. No presentan atributos específicos pero sí se les pueden asignar los atributos comunes de HTML básicos, de internacionalización y de eventos.



Cómo estructurar un texto

```
<html>
```

```
<head>
```

```
<title>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>
```

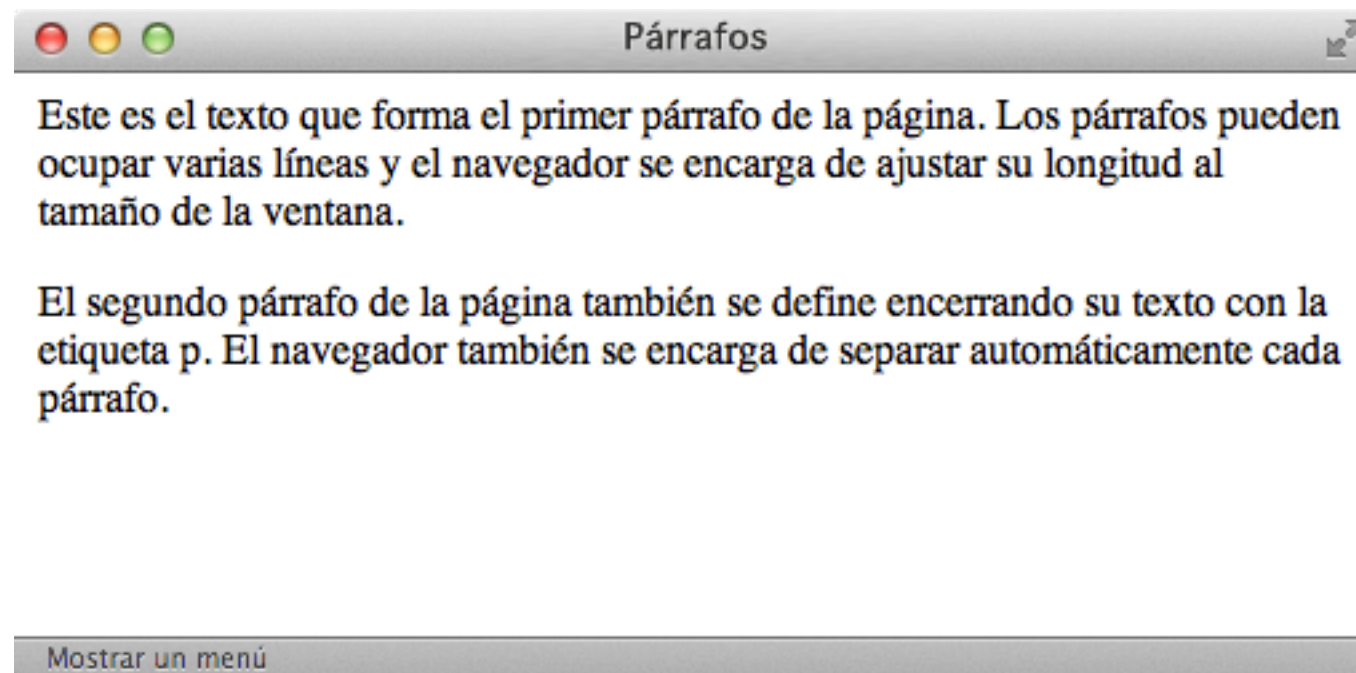
```
</body>
```

```
</html>
```

<http://cursosdedesarrollo.com/>



Cómo estructurar un texto



Cómo estructurar un texto

SECCIONES

Etiqueta <h1>,<h2>,<h3>,<h4>,<h5>,<h6>

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: -

Tipo de elemento: En bloque

Descripción: Define los títulos de las secciones de mayor importancia de la página



Cómo estructurar un texto

Estas etiquetas `<h1>...<h6>` definen títulos de sección, no secciones completas. Como se puede ver en la tabla anterior y al igual que el tag `<p>`, estas secciones son elementos en bloque, por lo que ocupan toda la anchura del navegador y tampoco presentan atributos específicos pero sí se les pueden asignar los atributos comunes de HTML básicos, de internacionalización y de eventos.

Los navegadores asignan de forma automática el tamaño del título de cada sección en función de su importancia, que se puede modificar utilizando las hojas de estilos CSS.



Cómo estructurar un texto

```
<html>
```

```
  <head>
```

```
    <title>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>
```

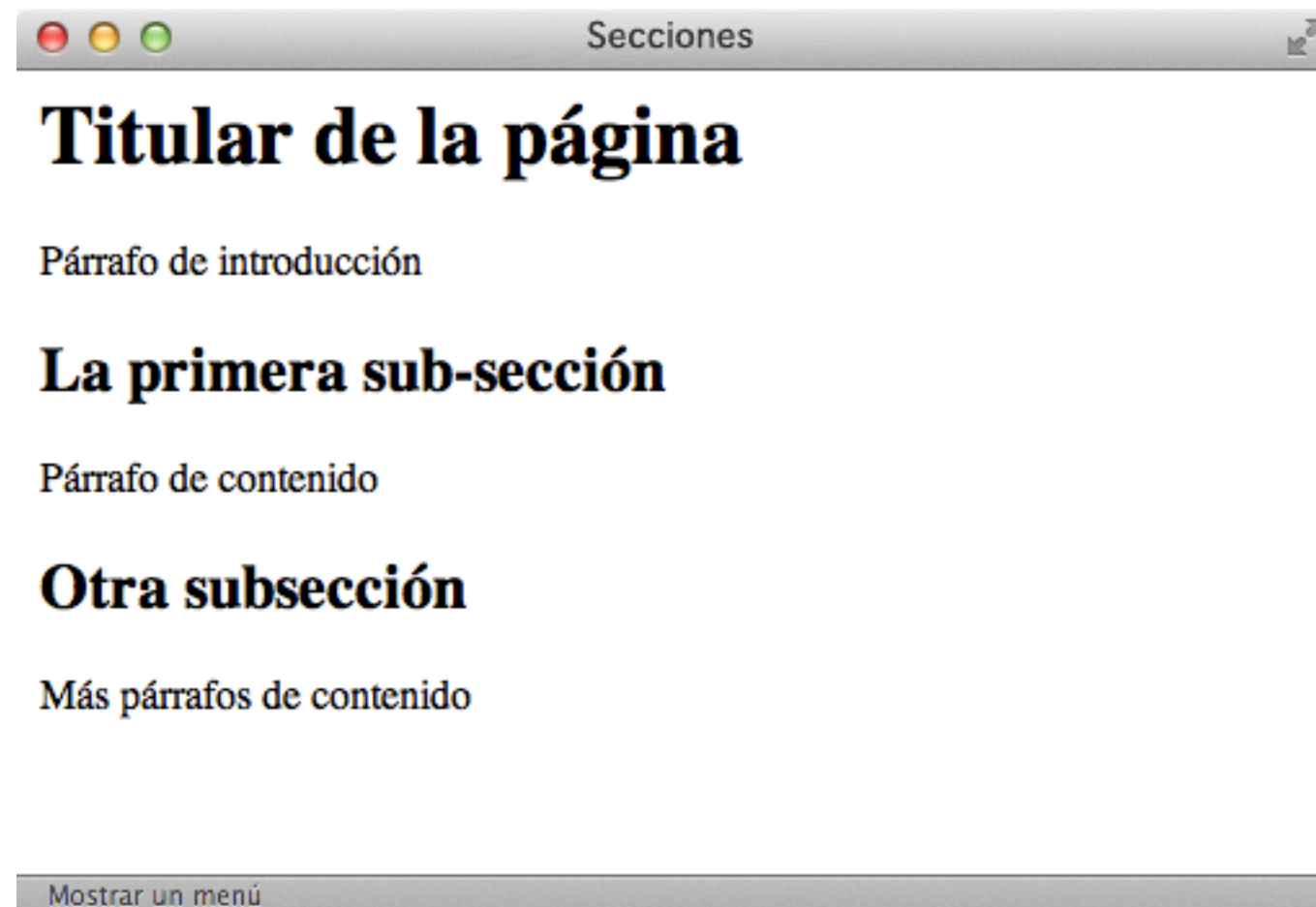
```
  </body>
```

```
</html>
```

<http://cursosdedesarrollo.com/>



Cómo estructurar un texto



Cómo estructurar un texto

MARCADO BÁSICO DE UN TEXTO

El siguiente paso posterior a la estructuración del texto consistiría en el marcado de éste. Para el marcado básico de un texto encontraríamos estas etiquetas



Cómo estructurar un texto

Etiqueta

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: -

Tipo de elemento: En línea

Descripción: Realza la importancia del texto que encierra - cursiva



Cómo estructurar un texto

Etiqueta

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: -

Tipo de elemento: En línea

Descripción: Realza con la máxima importancia el texto que encierra - negrita



Cómo estructurar un texto

```
<html>
```

```
<head>
```

```
<title>Etiquetas 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>
```

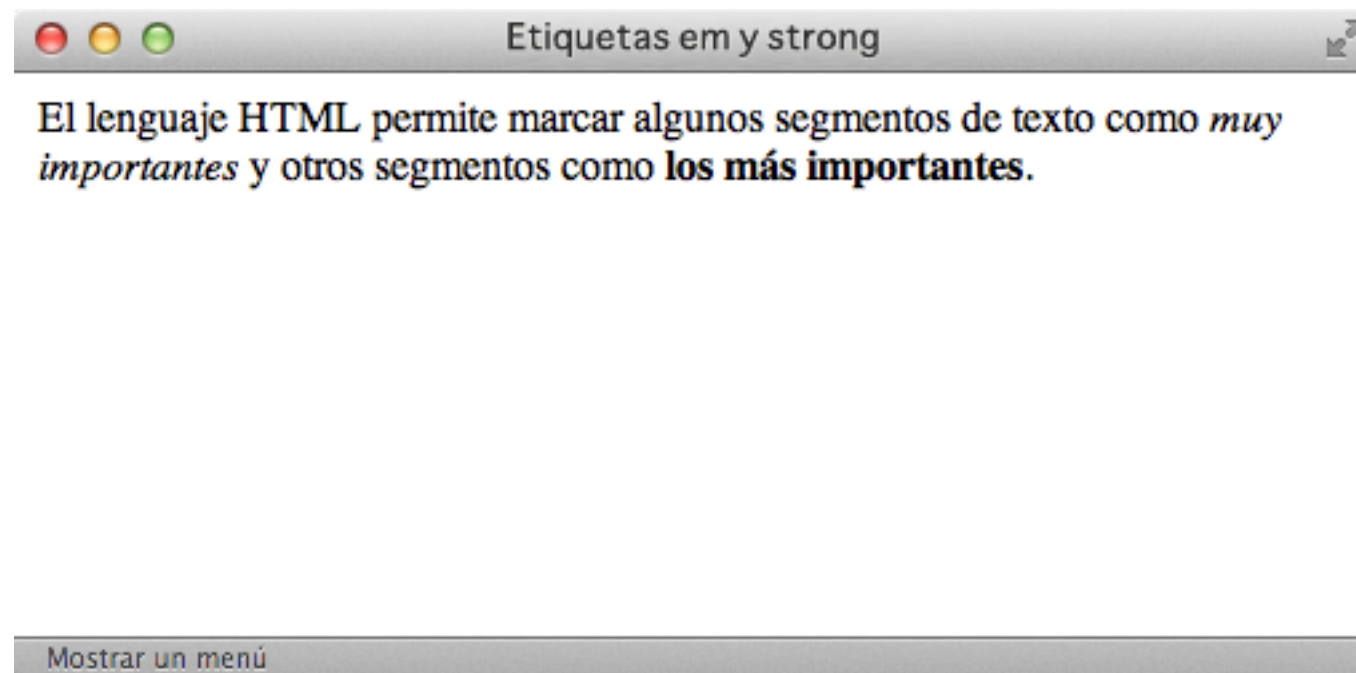
```
</body>
```

```
</html>
```

<http://cursosdedesarrollo.com/>



Cómo estructurar un texto



Cómo estructurar un texto

HTML también permite marcar de forma adecuada las modificaciones realizadas en el contenido de una página: el texto que ha sido eliminado y el texto que ha sido añadido



Cómo estructurar un texto

Etiqueta `<ins>`

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: `cite="url"` Indica la URL de la página en la que se puede obtener más información sobre el motivo por el que se realizó la modificación. `datetime="fecha"` Especifica la fecha y hora en la que se realizó el cambio

Tipo de elemento: En bloque y en línea

Descripción: Se emplea para marcar una modificación en los contenidos originales consistente en la inserción de un nuevo contenido



Cómo estructurar un texto

Etiqueta ``

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: `cite="url"` Indica la URL de la página en la que se puede obtener más información sobre el motivo por el que se realizó la modificación. `datetime="fecha"` Especifica la fecha y hora en la que se realizó el cambio

Tipo de elemento: En bloque y en línea

Descripción: Se emplea para marcar una modificación en los contenidos originales consistente en el borrado de cierto contenido



Cómo estructurar un texto

Por defecto, el texto eliminado () se muestra tachado y el texto insertado (<ins>) se muestra subrayado



Cómo estructurar un texto

```
<html>
```

```
<head>
```

```
<title>Etiquetas ins y del</title>
```

```
</head>
```

```
<body>
```

```
<p>El HTML, acrónimo inglés de Hyper Text Markup Language (lenguaje de <del  
datetime="20091025" cite="http://www.enlace.com">marcado de hipertexto</del> <ins  
datetime="20091026" cite="http://enlace.com">marcas hipertextuales</ins>) es un  
lenguaje de marcación diseñado para estructurar textos y presentarlos en forma de  
hipertexto.</p>
```

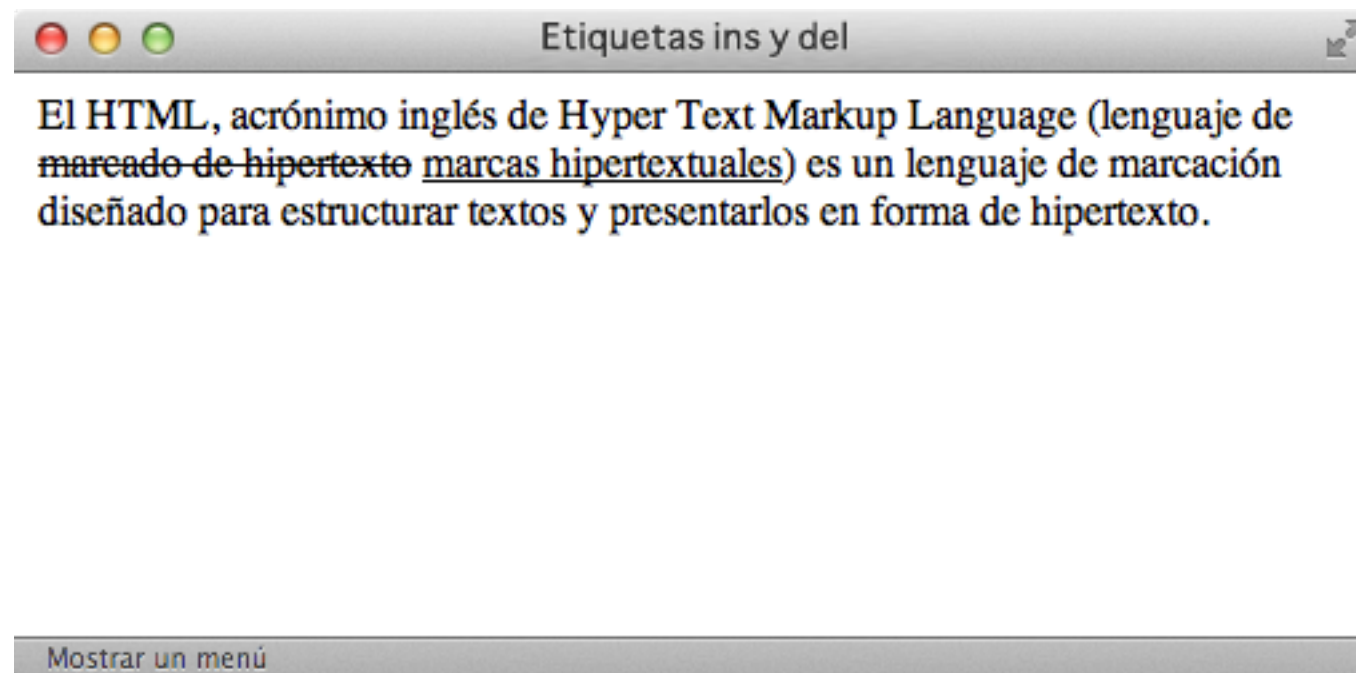
```
</body>
```

```
</html>
```

<http://cursosdedesarrollo.com/>



Cómo estructurar un texto



Cómo estructurar un texto

Además de estas dos, encontramos la etiqueta `<blockquote>` para incluir citas textuales en las páginas web.



Cómo estructurar un texto

Etiqueta `<blockquote>`

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: `cite="url"` Indica la URL de la página en la que se puede obtener más información sobre el motivo por el que se realizó la modificación

Tipo de elemento: En bloque

Descripción: Se emplea para indicar que el texto que encierra es una cita textual de otro texto externo



Cómo estructurar un texto

```
<html>
```

```
<head>
```

```
<title>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>
```

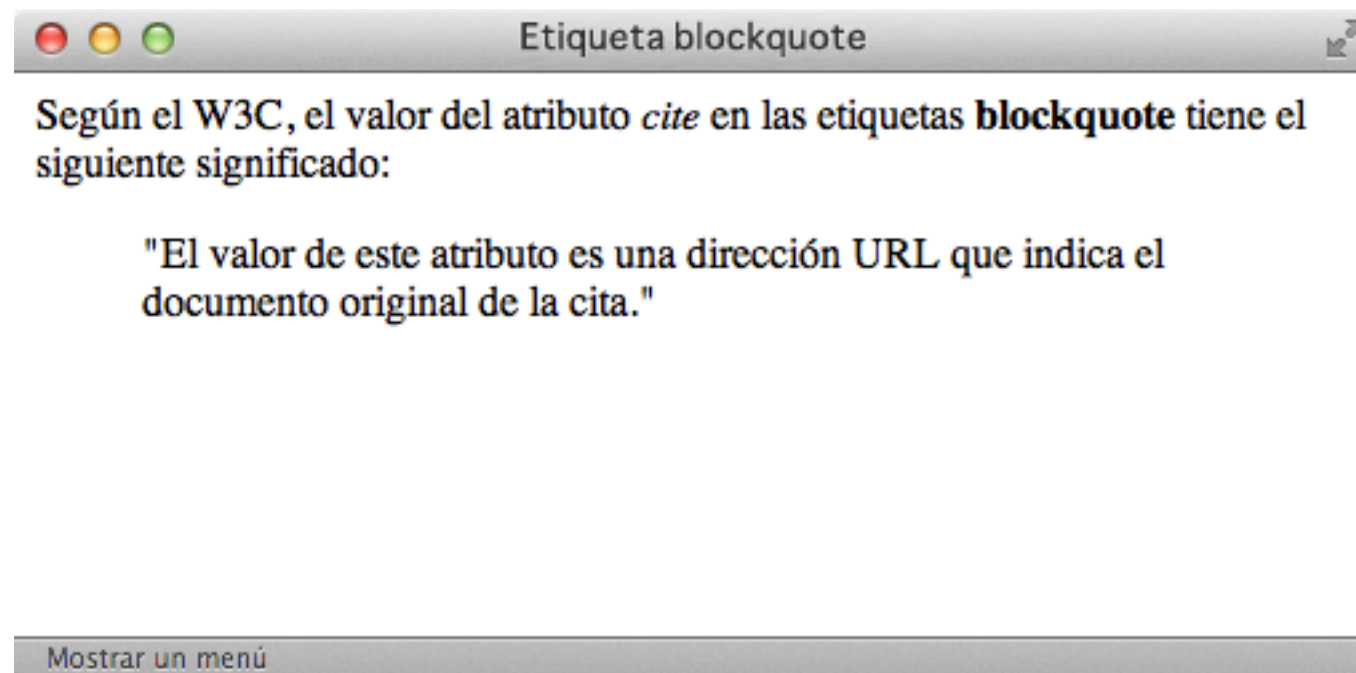
```
</body>
```

```
</html>
```

<http://cursosdedesarrollo.com/>



Cómo estructurar un texto



Marcado avanzado de un texto

La etiqueta `<abbr>` marca las abreviaturas de un texto y la etiqueta `<acronym>` se emplea para marcar las siglas o acrónimos del texto. En ambos casos, el atributo `title` se puede utilizar para incluir el significado completo de la abreviatura o sigla. La mayoría de navegadores muestran por defecto un borde inferior punteado para estos elementos.



Marcado avanzado de un texto

Etiqueta <abbr>

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: title="texto" Indica el significado completo de la abreviatura

Tipo de elemento: En línea

Descripción: Se emplea para marcar las abreviaturas del texto y proporcionar el significado de esas abreviaturas



Marcado avanzado de un texto

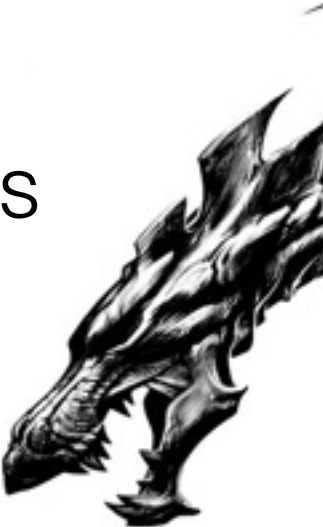
Etiqueta <acronym>

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: title="texto" Indica el significado completo del acrónimo o sigla

Tipo de elemento: En línea

Descripción: Se emplea para marcar las siglas o acrónimos del texto y proporcionar el significado de esas siglas



Marcado avanzado de un texto

```
<html>
```

```
<head>
```

```
<title>Etiqueta acronym</title>
```

```
</head>
```

```
<body>
```

```
<p>El lenguaje
```

```
<acronym title="HyperText Markup Language">HTML</acronym>
```

es estandarizado por el

```
<acronym title="World Wide Web Consortium">W3C</acronym>.</p>
```

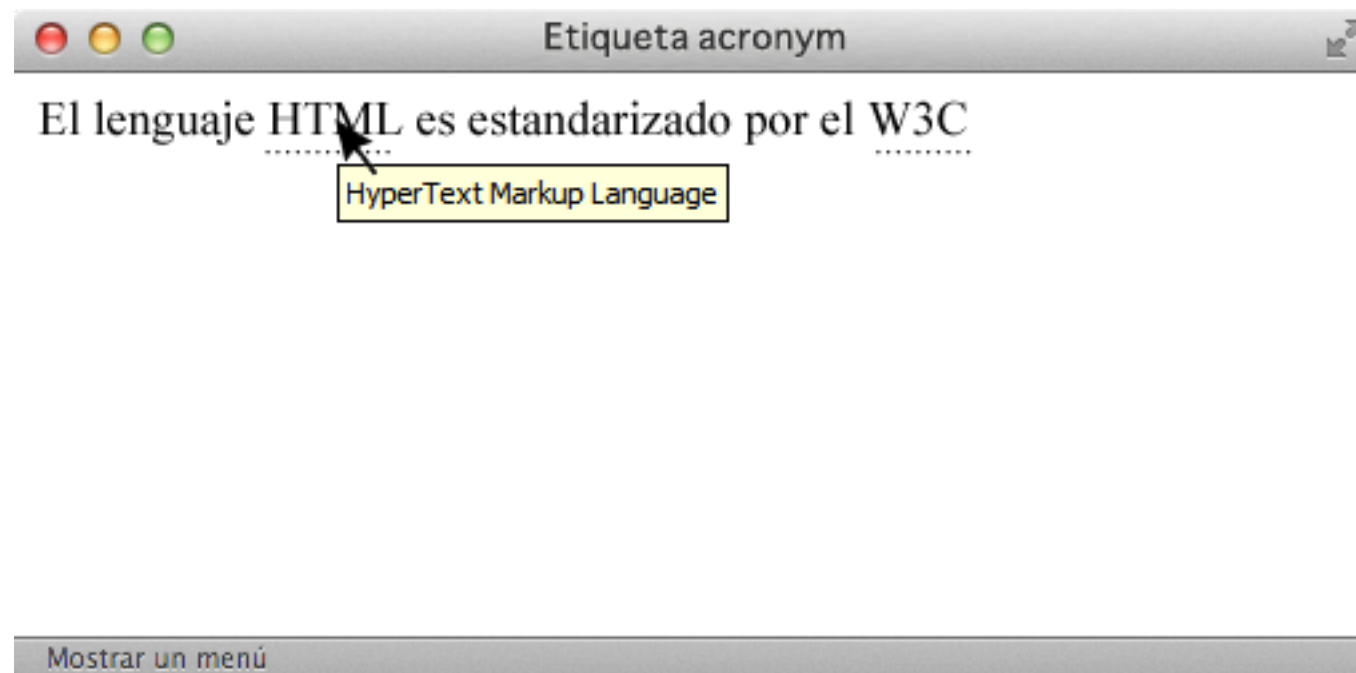
```
</body>
```

```
</html>
```

<http://cursosdedesarrollo.com/>



Marcado avanzado de un texto



Marcado avanzado de un texto

La etiqueta `<dfn>` proporciona al usuario la definición de todas las palabras para las que se considere apropiado, y la etiqueta `<cite>` se utiliza para marcar un texto como una citación



Marcado avanzado de un texto

Etiqueta `<dfn>`

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: `title="texto"` Indica el significado completo del término

Tipo de elemento: En línea

Descripción: Se emplea para marcar las definiciones de ciertos términos y proporcionar el significado de esos términos



Marcado avanzado de un texto

<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>



Marcado avanzado de un texto

Etiqueta <cite>

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: -

Tipo de elemento: En línea

Descripción: Se emplea para marcar una cita o una referencia a otras fuentes



Marcado avanzado de un texto

La diferencia entre `<cite>` y `<blockquote>` está en que el elemento `<cite>` marca el autor de la cita (persona, documento, etc.) y `<blockquote>` marca el contenido de la propia cita.



Marcado avanzado de un texto

Como dijo `<cite>Mahatma Gandhi</cite>`:

`<blockquote>Vive como si fueras a morir mañana y aprende como si fueras a vivir para siempre.</blockquote>`



Marcado avanzado de un texto

ETIQUETA

La etiqueta permite agrupar varios elementos en línea seguidos dentro de un mismo bloque (por ejemplo, varias palabras seguidas en un párrafo), para después darles formato con las hojas de estilo CSS. Se emplea para marcar cualquier elemento que no se puede marcar con las otras etiquetas definidas.



Marcado avanzado de un texto

La etiqueta `` se visualiza por defecto con el mismo aspecto que el texto normal. Por tanto, es habitual utilizar esta etiqueta junto con los atributos `id` y `class` para modificar posteriormente su aspecto con CSS.

Este tag sólo se puede utilizar, como ya hemos mencionado, para encerrar contenidos y etiquetas en línea. Cuando se quieren estructurar elementos en bloque, se utiliza la etiqueta `<div>`.



Marcado avanzado de un texto

Importante: Si quiere ponerse en contacto con nuestra empresa, puede hacerlo en el teléfono `0034 900 000 000` o a través de la dirección de correo `contacto@empresa.com`



Marcado avanzado de un texto

ESPACIOS EN BLANCO

El lenguaje HTML considera espacio en blanco a: los espacios en blanco, los tabuladores, los retornos de carro y el carácter de nueva línea. 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.



Marcado avanzado de un texto

```
<html>
```

```
<head>
```

```
<title>Espacios en blanco</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>
```

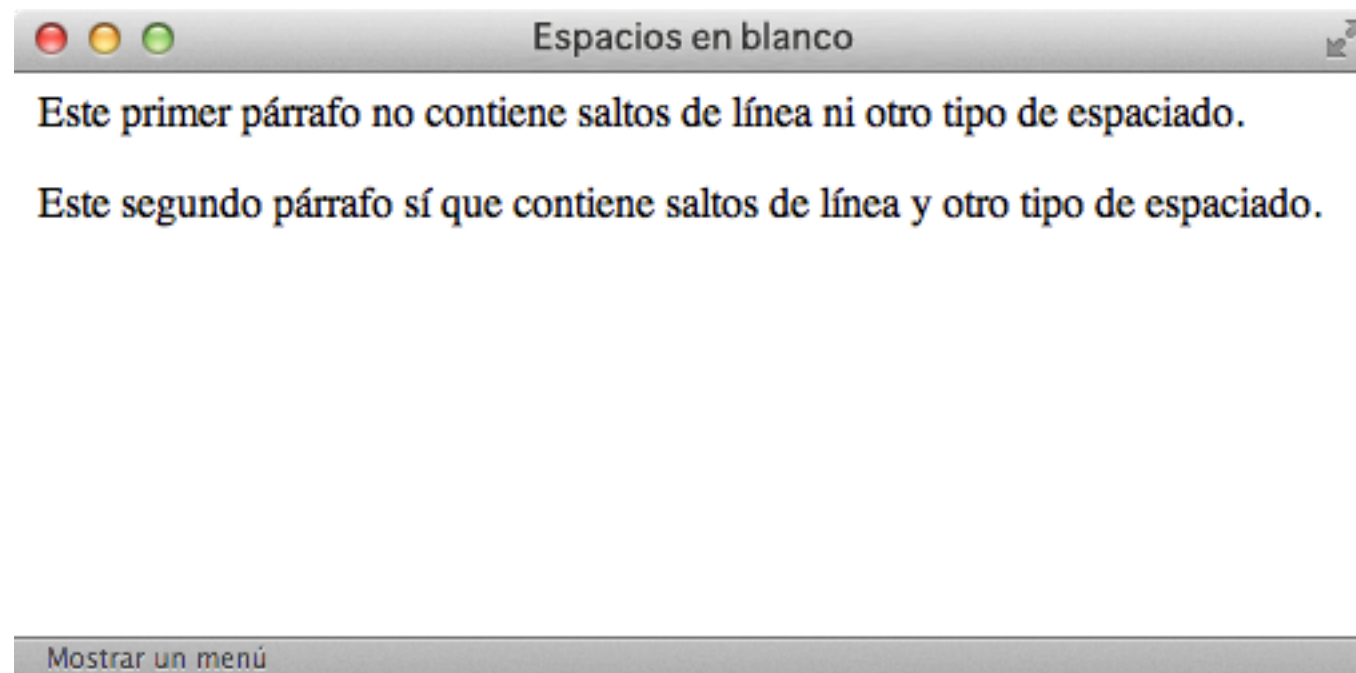
```
</body>
```

```
</html>
```

<http://cursosdedesarrollo.com/>



Marcado avanzado de un texto



Marcado avanzado de un texto

Sin embargo, existen alternativas para incluir espacios en blanco adicionales. Esto se consigue sustituyendo cada nuevo espacio en blanco por el texto ` `;

Cada texto ` ` equivale a un sólo espacio en blanco, por lo que se deben escribir tantos ` ` seguidos como espacios en blanco se quieran conseguir.



Marcado avanzado de un texto

```
<html>
```

```
<head>
```

```
<title>Espacios en blanco</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>
```

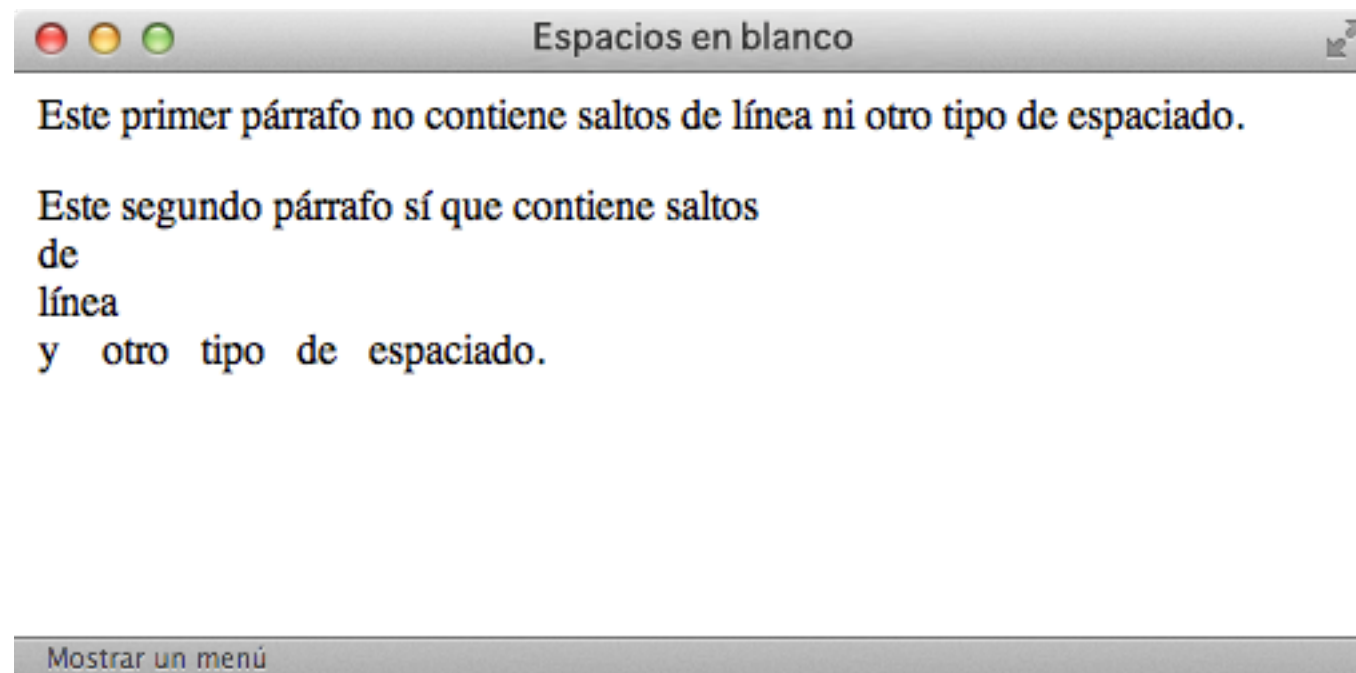
```
</body>
```

```
</html>
```

<http://cursosdedesarrollo.com/>



Marcado avanzado de un texto



Marcado avanzado de un texto

NUEVAS LÍNEAS

Para forzar una nueva línea, o lo que es lo mismo, lo equivalente a presionar la tecla Enter o Intro escribiendo un texto, se utiliza el tag `
`. Se trata de una etiqueta vacía, es decir, no encierra ningún texto.



Marcado avanzado de un texto

Etiqueta

Atributos comunes: básicos

Atributos propios: -

Tipo de elemento: En línea y etiqueta vacía

Descripción: Fuerza al navegador a insertar una nueva línea



Marcado avanzado de un texto

Existen varias formas de expresar esta etiqueta. Si bien se puede abrir y cerrar de forma consecutiva (`
</br>`), la forma de uso más común es abriendo y cerrando un único tag de esta forma: `
` o `
`.



Marcado avanzado de un texto

TEXTO PREFORMATEADO

Existe una manera de mostrar el texto tal y como está escrito, respetando los espacios en blanco y las nuevas líneas. Se utiliza, por ejemplo, cuando una página debe mostrar directamente el texto generado por alguna aplicación.



Marcado avanzado de un texto

Etiqueta <pre>

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: -

Tipo de elemento: En bloque

Descripción: Muestra el texto que encierra tal y como está escrito



Marcado avanzado de un texto

Los elementos `<pre>` son algo especiales, ya que los navegadores les aplican algunas reglas:

Mantienen todos los espacios en blanco (tabuladores, espacios y nuevas líneas)

Muestra el texto con un tipo de letra de ancho fijo (todas las letras de la misma anchura)

No se ajusta la longitud de las líneas (las líneas largas producen un scroll horizontal), lo que provoca que la anchura de la página sea superior a la anchura de la ventana del navegador.



Marcado avanzado de un texto

<pre>

La etiqueta pre

respeta los espacios en blanco

y

muestra el texto

tal y como

está escrito

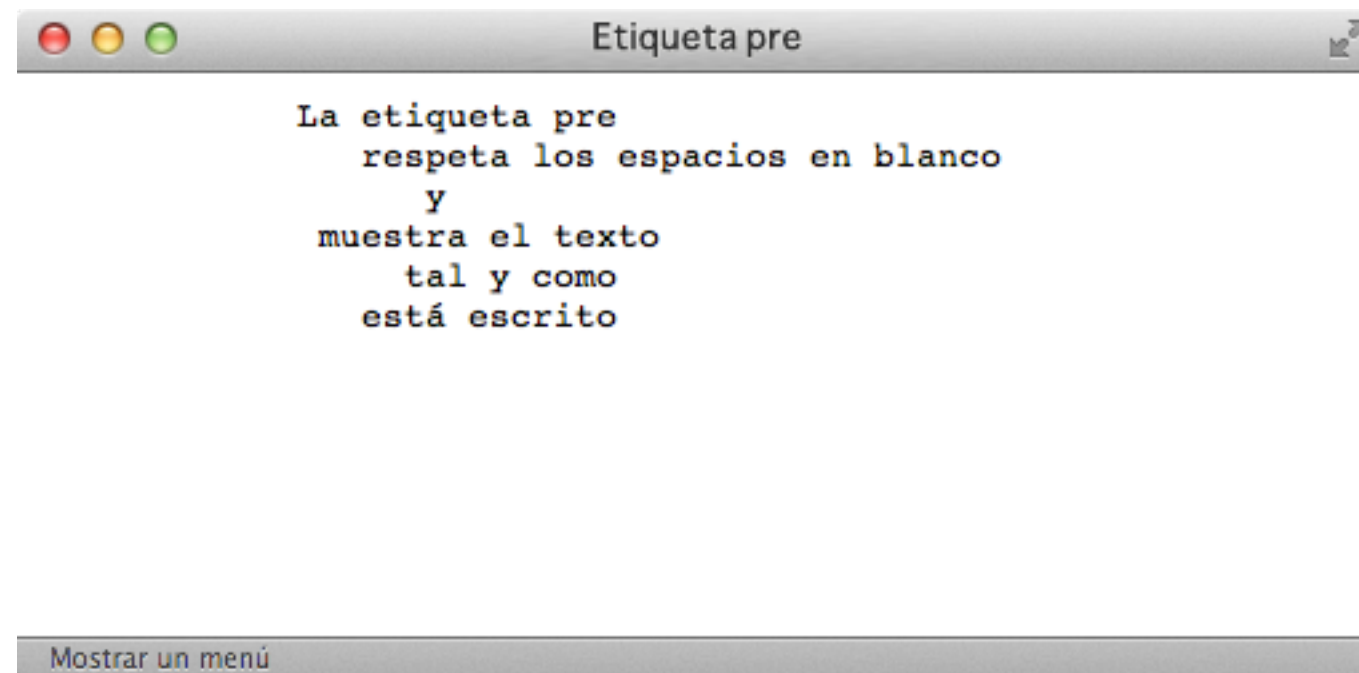
</p>

</pre>

<http://cursosdedesarrollo.com/>



Marcado avanzado de un texto



Marcado avanzado de un texto

También existe la etiqueta `<code>` que se utiliza para mostrar código fuente de cualquier lenguaje de programación.



Marcado avanzado de un texto

Etiqueta `<code>`

Atributos comunes: básicos, internacionalización, eventos

Atributos propios: -

Tipo de elemento: En bloque

Descripción: Delimita el texto considerado un fragmento de código fuente



Marcado avanzado de un texto

Las reglas que los navegadores aplican a las etiquetas `<code>` son:

Muestra el texto con un tipo de letra de ancho fijo (todas las letras de la misma anchura)

Sin embargo, no respeta los espacios en blanco ni las líneas, por lo que, en ese sentido, su comportamiento es similar a la etiqueta `<p>`.



Marcado avanzado de un texto

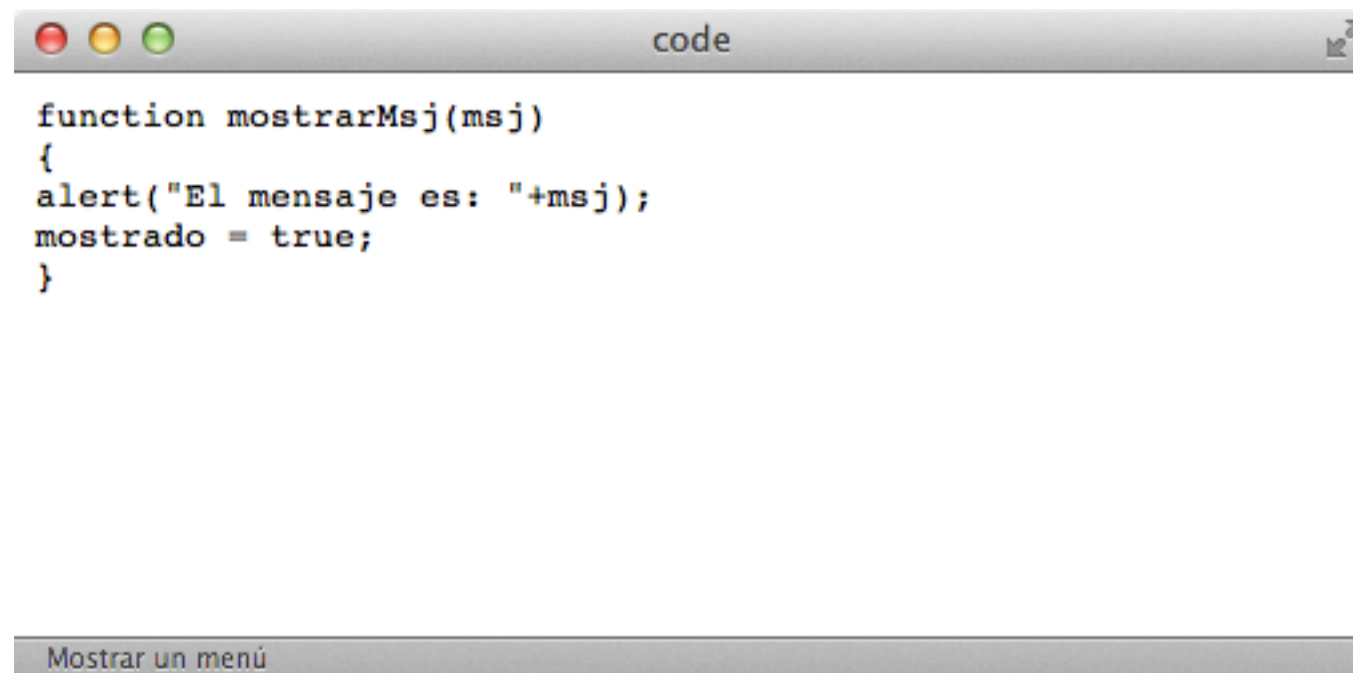
<code>

```
function mostrarMsj(msj)
{
    alert("El mensaje es: "+msj);
    mostrado = true;
}
```

</code>



Marcado avanzado de un texto



```
function mostrarMsj(msj)
{
  alert("El mensaje es: "+msj);
  mostrado = true;
}
```

Mostrar un menú



Codificación de caracteres

Existen caracteres que se utilizan habitualmente en los textos que no se pueden incluir directamente en las páginas web o que pueden darnos problemas. Estos son:

Los caracteres que utiliza el lenguaje HTML para definir sus etiquetas.

Los caracteres propios de idiomas que no son el inglés.



Codificación de caracteres

Para los caracteres propios del lenguaje HTML existen algunas expresiones o entidades HTML que los sustituyen:

Entidad	Caracter
---------	----------

<	<
------	---

>	>
------	---

&	&
-------	---

"	"
--------	---

 	espacio en blanco
--------	-------------------

'	'
--------	---



Codificación de caracteres

En cuanto a los caracteres propios de idiomas diferentes al inglés, cuando todos los procesos involucrados (entorno de desarrollo, servidor web y navegador) utilizan la misma codificación de caracteres (por ejemplo, UTF-8), el texto se verá correctamente en el navegador. Pero si la codificación cambia sin realizar una conversión correcta, el navegador mostrará caracteres extraños.



Codificación de caracteres

La solución más correcta y sencilla es sustituir estos caracteres potencialmente problemáticos por su entidad HTML, ya que si se utilizan las entidades HTML en vez de los caracteres problemáticos, es indiferente pasar de una codificación de caracteres a otra diferente.



Codificación de caracteres

Entidad Caracter

ñ ñ

Ñ Ñ

á á

é é

í í

ó ó

ú ú



Codificación de caracteres

Entidad	Caracter
---------	----------

Á	Á
----------	---

É	É
----------	---

Í	Í
----------	---

Ó	Ó
----------	---

Ú	Ú
----------	---

€	€
--------	---



Codificación de caracteres

Puedes consultar la tabla completa de entidades HTML en
Wikipedia: [http://en.wikipedia.org/wiki/
List_of_XML_and_HTML_character_entity_references](http://en.wikipedia.org/wiki/List_of_XML_and_HTML_character_entity_references)



Codificación de caracteres

De cara a marcar la utilización de una determinada codificación de caracteres en el fichero html dentro de la etiqueta <head> hay que introducir la etiqueta <meta>

```
<meta http-equiv="Content-Type" content="text/html;  
      charset=UTF-8" />
```



Conclusiones

Mediante el uso de
etiquetas podemos
maquetar el texto y
marcarlo a nivel semántico



Datos de Contacto

<http://www.cursosdedesarrollo.com>
info@cursosdedesarrollo.com

<http://cursosdedesarrollo.com/>



Licencia



David Vaquero Santiago

Esta obra está bajo una
Licencia Creative Commons Atribución-
NoComercial-CompartirIgual 4.0
Internacional

Derivada de:

<http://www.arkaitzgarro.com/xhtml/>

y

<http://javiereguiluz.com/>

<http://cursosdedesarrollo.com/>

