

Sintaxis del HTML

La sintaxis de un lenguaje es el "conjunto de reglas que definen las secuencias correctas de los elementos de un lenguaje de programación" ([diccionario RAE](#)). Aunque el HTML no es un lenguaje de programación, sino un lenguaje de marcas, en el HTML también existen una serie de reglas para que los documentos sean correctos. En general, la sintaxis del HTML está inspirada en la [norma SGML](#) (aunque no la cumple estrictamente).

Si un documento contiene errores de sintaxis se dice que es un documento **inválido**. Aunque un documento HTML no sea sintácticamente correcto, los navegadores intentan mostrar el documento de la mejor manera posible. Aunque los navegadores a menudo consiguen mostrar correctamente documentos inválidos, se aconseja validar y corregir los documentos para asegurar que los navegadores puedan mostrarlos correctamente.

Nota: En HTML 5 se define también una sintaxis llamada XHTML que está basada en el XML (aunque tampoco la cumple estrictamente). En general, la sintaxis XHTML es más "estricta", en el sentido de imponer restricciones en la forma de escribir etiquetas, atributos o valores, mientras que la sintaxis HTML es más "flexible". las diferencias entre ambas sintaxis se comenta en la lección [Diferencias entre HTML y XHTML](#).

Generalidades

- El carácter **<** (menor que) sólo puede utilizarse para indicar el comienzo de una etiqueta. Si se quiere mostrar el carácter **<**, se debe escribir la entidad de carácter **<**.



```
<p> 2 &lt; 5 </p>
```



```
<p>2 < 5</p>
```

- Los nombres de los elementos, los nombres de los atributos y los valores de los atributos se suelen escribir en minúsculas, aunque se pueden escribir en mayúsculas.



```
<p class="xxx">aaaa</p>
```



```
<P CLASS="XXX">aaaa</P>
```

- Las entidades numéricas expresadas en hexadecimal se pueden escribir en minúsculas o mayúsculas.



```
<p>&#x3b1; &#x3b2; &#x3c0;</p>
```



```
<p>&#X3b1; &#x3B2; &#X3C0;</p>
```

Elementos (etiquetas)

- El nombre del elemento debe ir a continuación del carácter **<** (sin espacios).



```
<p>AAAA</p>
```



```
< p>AAAA</p>
```

- El nombre del elemento en la etiqueta de cierre tiene que coincidir con el nombre en la etiqueta de apertura.



```
<p>AAAA</p>
```



```
<p>AAAA</h2>
```

- Los documentos HTML suelen tener un único elemento raíz de todo el documento, el elemento `<html>`, aunque puede omitirse.



```
<!DOCTYPE html>
<html>
<head>
  <meta charset="utf-8">
  <title>HTML</title>
</head>
<body>
  <!-- contenido de la página -->
</body>
</html>
```



```
<!DOCTYPE html>
<head>
  <meta charset="utf-8">
  <title>HTML</title>
</head>
<body>
  <!-- contenido de la página -->
</body>
```

- Todos los elementos tienen que estar completamente contenidos en otros elementos (no se pueden cruzar las etiquetas).



```
<p>AAAA <strong>BBBB <em>CCCC</em> </strong> <em>DDDD</em> EEEE</p>
```



```
<p>AAAA <strong>BBBB <em>CCCC</strong> </em> <em>DDDD</em> EEEE</p>
```

- Los elementos no vacíos (los que pueden contener texto) suelen tener etiquetas de cierre, aunque la norma define algunas circunstancias en las se pueden omitir las etiquetas de cierre.



```
<p>AAAA</p>
<p>BBBB</p>
```



```
<p>AAAA
```

```
<p>BBBB</p>
```

- Los elementos vacíos (los que no pueden contener texto) no tienen etiqueta de cierre. La etiqueta de apertura puede acabar en `" />"`.

Las etiquetas vacías (*void*) de HTML son `<area>`, `<base>`, `
`, `<col>`, `<embed>`, `<hr>`, ``, `<input>`, `<link>`, `<meta>`, `<param>`, `<source>`, `<track>` y `<wbr>`.



```
<hr>
```



```
<hr />
```



```
<hr></hr>
```

- Algunos elementos no pueden contener otros elementos, aunque no he visto en la recomendación un apartado dedicado a este tema.
 - El elemento `<a>` no puede contener elementos `<a>`.



```
<p><a href="https://www.mclibre.org"><a href="https://www.mclibre.org">AAAA</a></a></p>
```

- El elemento `<button>` no puede contener elementos `<button>`, `<fieldset>`, `<form>`, `<input>`, `<label>`, `<select>` o `<textarea>`.
- El elemento `<form>` no puede contener elementos `<form>`.



```
<form action="prueba.php"><form action="prueba.php"></form></form>
```

- El elemento `<label>` no puede contener elementos `<label>`.



```
<p><label><label><input type="checkbox" name="casilla">Casilla 1</label></label></p>
```

Atributos

- Los atributos sólo pueden aparecer en la etiqueta de apertura:



```
<p class="xxx yyy">AAAA</p>
```



```
<p>AAAA</p class="xxx">
```

- Un atributo sólo puede aparecer una vez en cada etiqueta de apertura o en cada etiqueta vacía (es decir, en una etiqueta los atributos no pueden repetirse).



```
<p class="xxx yyy">AAAA</p>
```



```
<p class="xxx" class="yyy">AAAA</p>
```

- Los valores de los atributos se suelen escribir entre comillas (dobles o simples), pero no es necesario si el valor no contiene espacios en blanco.

```
<p style="color: red">AAAA</p>
```



```
<p style=color:red>AAAA</p>
```



```
<p style=color: red>AAAA</p>
```



```
<p class="xxx">AAAA</p>
```



```
<p class=xxx>AAAA</p>
```

- Los espacios en blanco iniciales y finales en los valores de los atributos no se tienen en cuenta. En los valores de los atributos, los saltos de línea o si hay varios espacios en blanco seguidos se consideran como un único espacio en blanco.



```
<p class="xxx">AAAA</p>
```



```
<p class="   xxx   ">AAAA</p>
```

- Los atributos booleanos (los que sólo pueden tener un valor) pueden no tener valor asignado y se interpreta entonces como si se hubiera establecido ese único valor posible.



```
<ol reversed>
  <li>AAA</li>
  <li>BBB</li>
</ol>
```



```
<ol reversed="reversed">
  <li>AAA</li>
  <li>BBB</li>
</ol>
```



```
<ol reversed="">
  <li>AAA</li>
  <li>BBB</li>
</ol>
```

- El valor de un atributo **id** no se puede repetir en un documento (es decir, en un documento no puede haber dos elementos cualesquiera con el mismo valor en el atributo **id**).



```
<p id="xxx">AAAA</p>
<p id="yyy">BBBB</p>
```



```
<p id="xxx">AAAA</p>
<p id="xxx">BBBB</p>
```



```
<p id="xxx">AAAA</p>
<p><strong id="xxx">BBBB</strong></p>
```

- Los valores de atributos no pueden contener referencias a entidades externas.

Por completar

- El texto de sustitución de una entidad incluida en un atributo no puede contener el carácter "<".

Por completar

Última modificación de esta página: 5 de octubre de 2018



Esta página forma parte del curso [Páginas web HTML y hojas de estilo CSS](#) por [Bartolomé Sintes Marco](#) que se distribuye bajo una [Licencia Creative Commons Reconocimiento-CompartirIgual 4.0 Internacional \(CC BY-SA 4.0\)](#).