

Lenguajes de marcas y Sistemas de gestión de la información

HTML: Estructura del documento

¿QUÉ ES HTML?

HTML o **Híper Text Markup Language** es el lenguaje de marcas de texto más utilizado en la World Wide Web.

Fue creado en 1989 por Tim Berners-Lee como una particularización de SGML.

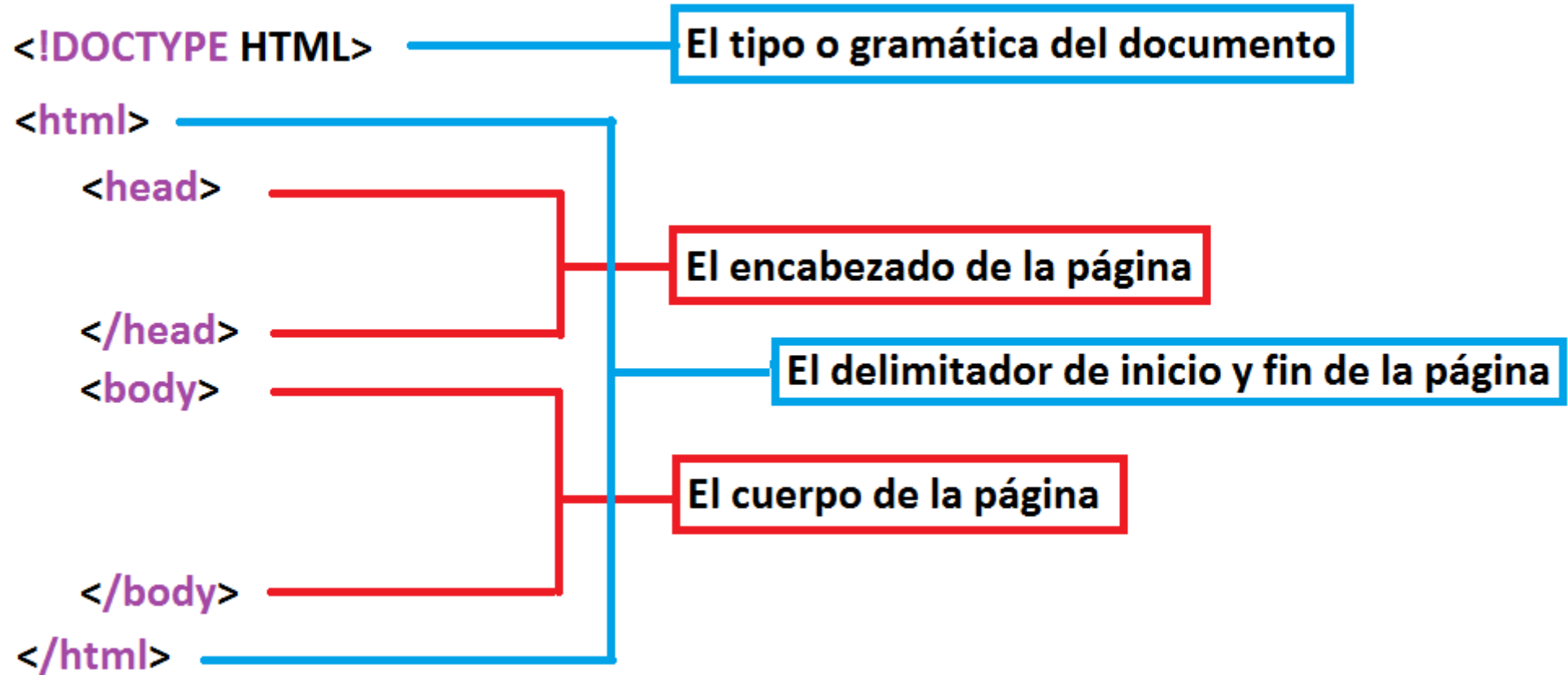
El entorno de desarrollo para el lenguaje HTML es un simple procesador de texto.

Se trata de un lenguaje interpretado en el navegador, de modo que los errores se detectarán en ejecución.

El lenguaje HTML permite organizar el texto y los objetos de una página, crear listas y tablas, definir hipervínculos, etc.

Los **hipervínculos** son enlaces desde una página web a otra página o a otro archivo.

ESTRUCTURA DEL DOCUMENTO



GRAMÁTICAS

Las **gramáticas** o **DTD** para HTML dependen de la especificación del W3C que utilicemos en nuestras páginas.

La especificación **HTML 4** contempla tres DTD:

- **DTD estricta** (HTML 4.0 Strict DTD), que incluye todos los elementos y atributos que no han sido desaprobados o deprecated.
- **DTD transicional o flexible** (HTML 4.0 Transitional DTD), que incluye todo lo anterior más los elementos y atributos desaprobados.
- **DTD para documentos con marcos** (HTML 4.0 Frameset DTD), que engloba la transicional más lo relativo a la creación de documentos con marcos o frames.

La especificación **HTML 5** no contempla ningún DTD, pues no está basada en SGML.

ETIQUETA <!DOCTYPE>

La etiqueta <!DOCTYPE> es obligatoria en cualquier documento HTML y tiene que aparecer en la primera línea del documento.

Define la especificación HTML utilizada en el documento del siguiente modo:

- **HTML 5:** <!DOCTYPE html>
- **HTML 4.01 Strict:**
 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
- **HTML 4.01 Transitional:**
 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">
- **HTML 4.01 Frameset:**
 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN" "http://www.w3.org/TR/html4/frameset.dtd">

ETIQUETA <html>

```
1 <!DOCTYPE html>
2 <html lang="es">
3     <head>
4
5     </head>
6     <body>
7
8     </body>
9 </html>
```

La etiqueta <html> es obligatoria e indica el inicio y fin del documento HTML.

El atributo **lang** define el lenguaje de la página.

ETIQUETA <head>

La etiqueta <head> es obligatoria y contiene los metadatos de la página.

Los **metadatos** son información general relacionada con la página y/o el sitio web, que no se visualizan en el navegador, pero afectan a su funcionamiento.

Las **etiquetas** que puede contener son: <title>, <meta>, <base> y <link>.

ETIQUETA `<body>`

La etiqueta `<body>` es obligatoria y contiene los elementos que componen la página y serán visualizados por el navegador.

Solo puede existir un elemento `<body>` dentro de cada documento.

Sintaxis HTML5

HTML5 es muy permisivo a la hora de escribir código, pues ofrece una gran flexibilidad y pocas restricciones, al contrario que sus antepasados:

- Las comillas para los valores de los atributos son opcionales y pueden utilizarse tanto dobles como simples.
- Los elementos HTML que no tienen etiqueta de cierre, como 'br' o 'img', ya no tienen que indicar el cierre del elemento con /.
- Permite el uso indistinto de mayúsculas y minúsculas en la sintaxis de etiquetas y atributos.

Sintaxis HTML5 – Reglas de estilo

Dicha permisividad puede generar un problema en un código en el que hayan trabajado varios desarrolladores y cada uno haya escrito el código a su manera, por lo que se recomienda seguir las siguientes pautas:

- Usar solo minúsculas.
- Usar siempre comillas dobles para los valores de los atributos.
- Cerrar con / las etiquetas sin cierre.
- Utilizar el sangrado para que el código sea más legible.