



# **HTML**

## **Capítulo 2**

*Lenguajes de Marcas y Sistemas de Gestión de Información*

*Curso 2019/2020*

# + Capítulo 2: CARÁCTERÍSTICAS BÁSICAS

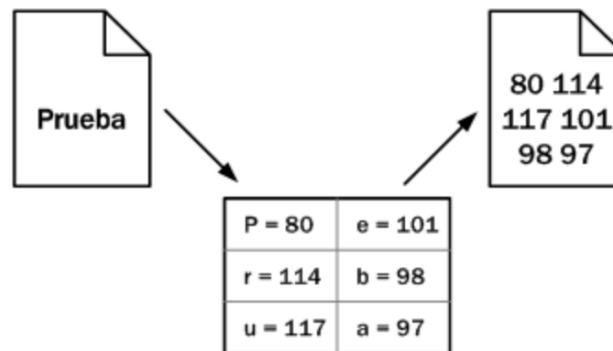
1. Lenguajes de Etiquetas
2. El primer documento HTML
3. Etiquetas y Atributos
4. Elementos HTML
5. Sintaxis de las etiquetas XHTML

# + 1. LENGUAJES DE ETIQUETAS

**RETO:** ¿Cómo almacenar la información en los archivos digitales?

Como los primeros archivos sólo contenían texto sin formato, la solución utilizada era muy sencilla: se **codificaban** las letras del alfabeto y se transformaban en números.

Para almacenar un contenido de texto en un archivo electrónico, se utiliza una **tabla de conversión** que transforma cada carácter en un número. Una vez almacenada la secuencia de números, el contenido del archivo se puede recuperar realizando el proceso inverso.





# 1. LENGUAJES DE ETIQUETAS

El proceso de transformación de caracteres en secuencias de números se denomina **codificación de caracteres** y cada una de las tablas que se han definido para realizar la transformación se conocen con el nombre de **páginas de código**. Una de las codificaciones más conocidas es la codificación ASCII.

Una vez resuelto el problema de almacenar el texto simple, se presenta el reto de almacenar los **contenidos de texto con formato**: *¿cómo se almacena un texto en negrita? ¿y un texto de color rojo? ¿y otro texto azul, en negrita y subrayado?*



# 1. LENGUAJES DE ETIQUETAS

Utilizar una tabla de conversión no es posible, ya que existen **infinitos** posibles **estilos** para aplicar al texto.

**SOLUCIÓN:** El archivo electrónico almacena tanto los contenidos como la información sobre el formato de esos contenidos.

*Ejemplo:* Se quiere dividir un texto en párrafos y se desea dar especial importancia a algunas palabras:

```
<parrafo>
```

```
Contenido de texto con <importante>algunas palabras</  
importante> resaltadas de forma especial.
```

```
</parrafo>
```



# 1. LENGUAJES DE ETIQUETAS

El proceso de indicar las diferentes partes que componen la información se denomina **marcar** (*markup*). Cada una de las palabras que se emplean para marcar el inicio y el final de una sección se denominan **etiquetas**.

Las etiquetas se indican por pares y se forman de la siguiente manera:

- **Etiqueta de Apertura:** carácter <, seguido del nombre de la etiqueta (sin espacios en blanco) y terminado con el carácter >.
- **Etiqueta de Cierre:** carácter <, seguido del carácter /, seguido del nombre de la etiqueta (sin espacios en blanco) y terminado con el carácter >.

La estructura típica de las etiquetas HTML es:

```
<nombre_etiqueta> ... </nombre_etiqueta>
```

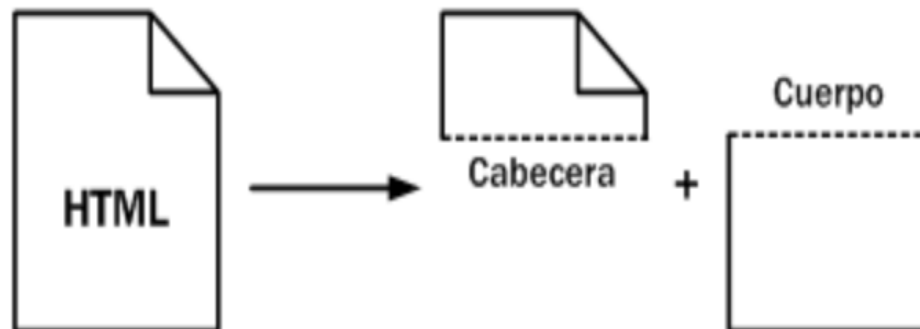


## 2. EL PRIMER DOCUMENTO HTML

Las páginas HTML se dividen en **dos partes**: la cabecera y el cuerpo.

La cabecera (*head*) incluye información sobre la propia página, como por ejemplo su título y su idioma.

El cuerpo (*body*) de la página incluye todos sus contenidos, como párrafos de texto e imágenes.



## + 2. EL PRIMER DOCUMENTO HTML

Código HTML de una página web muy sencilla:

```
<html>

  <head>

    <title>El primer documento HTML</title>

  </head>

  <body>

    <p>El lenguaje HTML es <strong>tan sencillo</strong> que prácticamente se entiende sin estudiar el significado de sus etiquetas principales.</p>

  </body>

</html>
```





## 2. EL PRIMER DOCUMENTO HTML

Es imprescindible que utilices un editor de texto **sin formato**. Si usas Windows, puedes utilizar el Bloc de Notas, Wordpad, Visual Studio Code, Sublime, Notepad++, etc... pero no puedes utilizar un procesador de textos como Word o Open Office.



Si utilizas sistemas operativos tipo Linux, puedes utilizar editores como Gedit, Kedit, Kate e incluso Vi, pero no utilices KOffice ni Open Office.

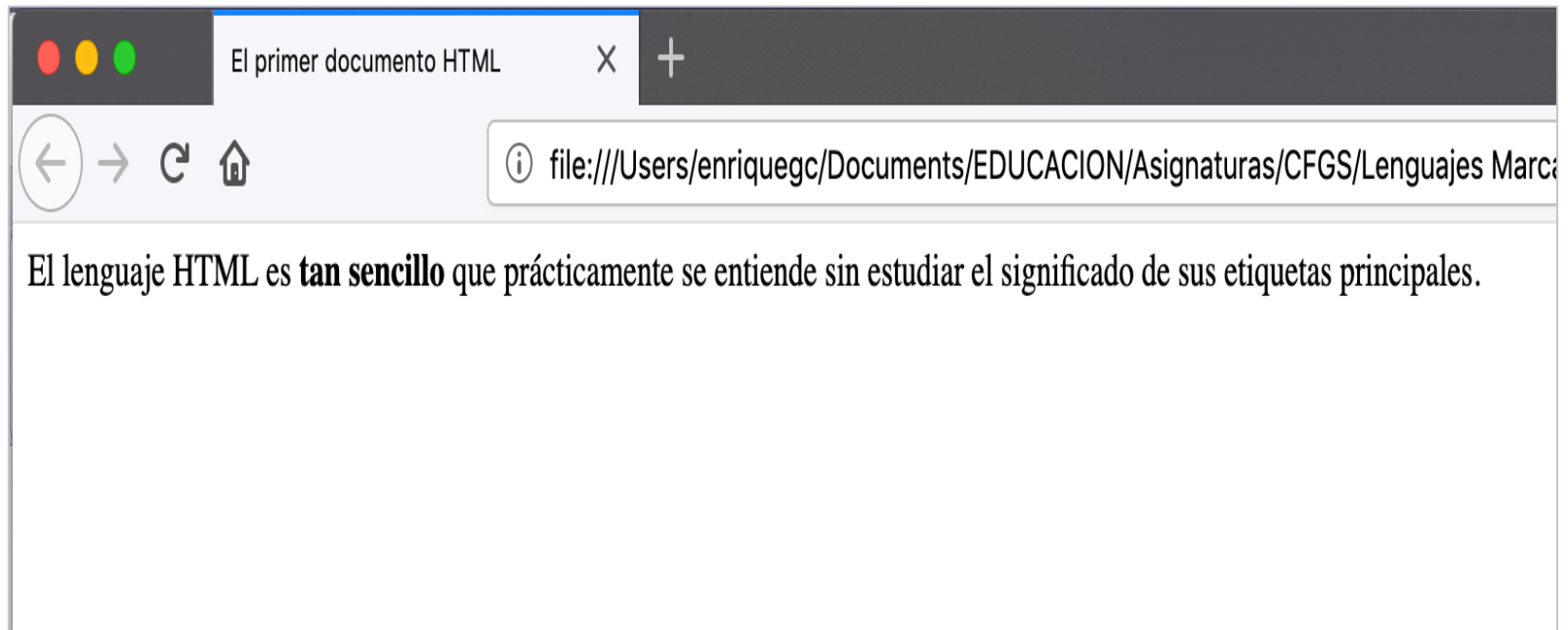




## 2. EL PRIMER DOCUMENTO HTML

10

Después de crear el archivo con el contenido HTML, ya se puede abrir con cualquier navegador para que se muestre con el siguiente aspecto:



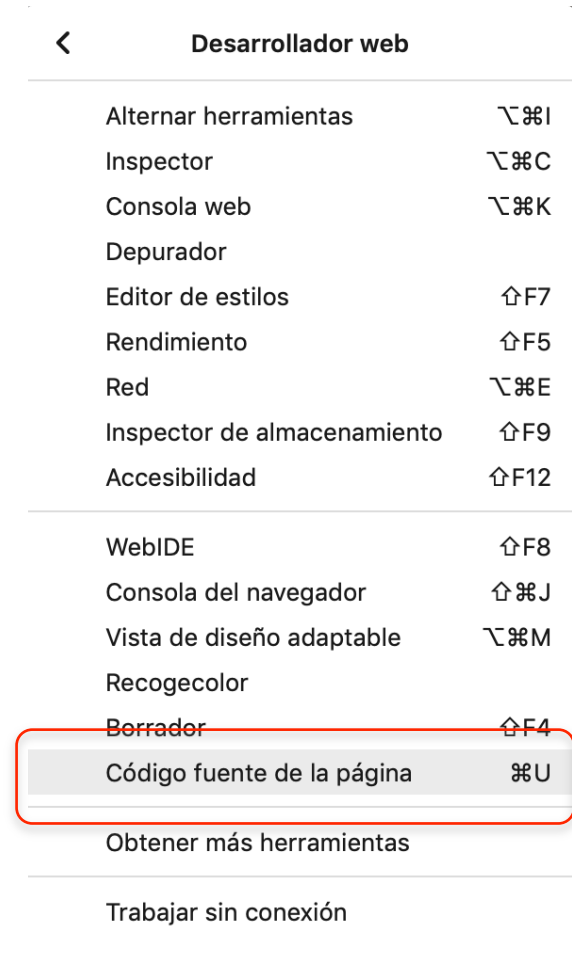


## 2. EL PRIMER DOCUMENTO HTML

11

Pulsa sobre el menú *Ver Código fuente* y podrás ver el código HTML de la página que está cargada en el navegador.

Volviendo al código HTML del primer ejemplo, es importante conocer las **tres etiquetas** principales de un documento HTML (`<html>`, `<head>`, `<body>`).

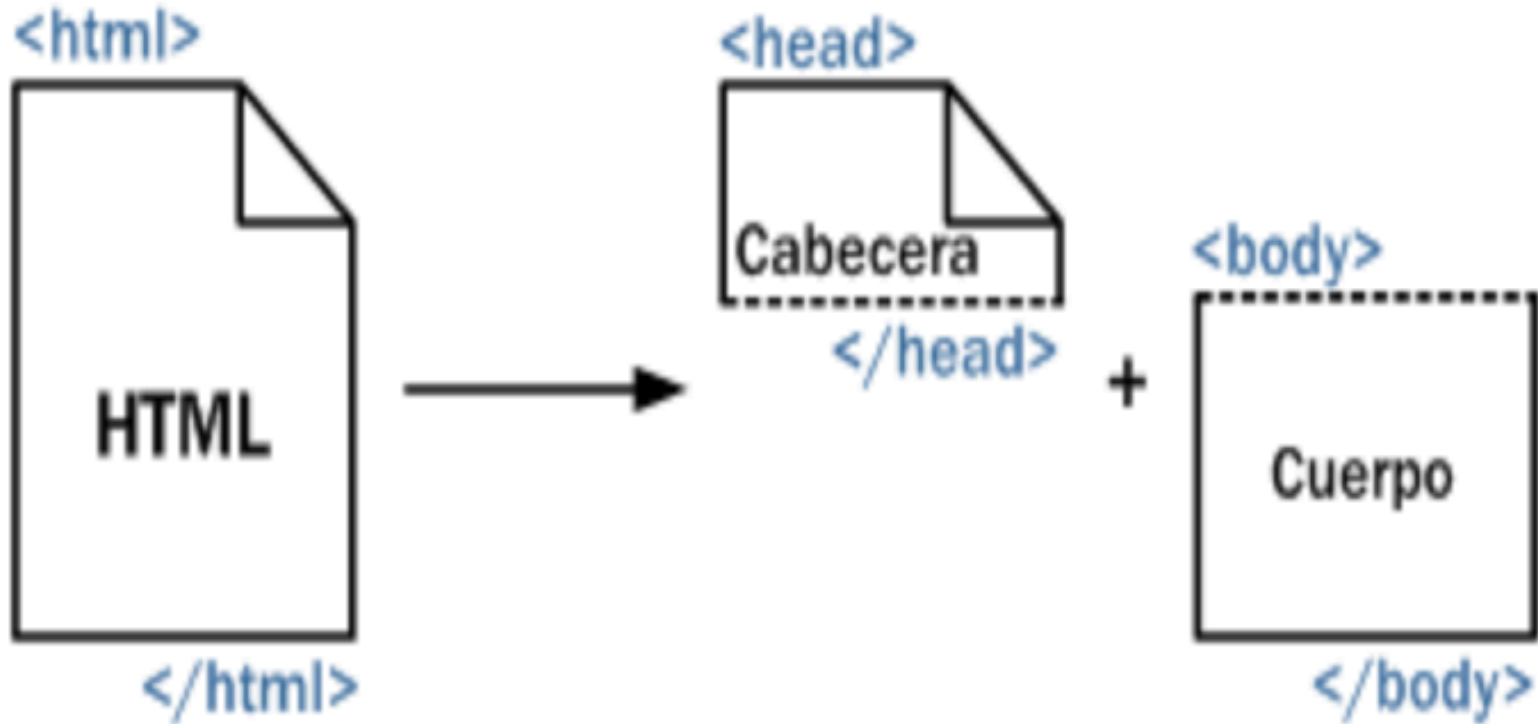




## 2. EL PRIMER DOCUMENTO HTML

- `<html>`: Indica el comienzo y el final de un documento HTML. Ninguna etiqueta o contenido puede colocarse antes o después de la etiqueta `<html>`. En el interior de la etiqueta `<html>` se definen la cabecera y el cuerpo del documento HTML y todo lo que se coloque fuera de la etiqueta `<html>` se ignora.
- `<head>`: Delimita la parte de la cabecera del documento y contiene información sobre el propio documento HTML, como por ejemplo su título y el idioma de la página. Los contenidos indicados en la cabecera no son visibles para el usuario, con la excepción de la etiqueta `<title>`, que se utiliza para indicar el título del documento y que los navegadores lo visualizan en la parte superior izquierda de la ventana del navegador.
- `<body>`: Delimita el cuerpo del documento HTML. El cuerpo encierra todos los contenidos que se muestran al usuario (párrafos de texto, imágenes, tablas).

## + 2. EL PRIMER DOCUMENTO HTML



## 3. ETIQUETAS Y ATRIBUTOS

HTML define **91 etiquetas** que los diseñadores pueden utilizar para marcar los diferentes elementos que componen una página:

a, abbr, acronym, address, applet, area, b, base, basefont, bdo, big, blockquote, body, br, button, caption, center, cite, code, col, colgroup, dd, del, dfn, dir, div, dl, dt, em, fieldset, font, form, frame, frameset, h1, h2, h3, h4, h5, h6, head, hr, html, i, iframe, img, input, ins, isindex, kbd, label, legend, li, link, map, menu, meta, noframes, noscript, object, ol, optgroup, option, p, param, pre, q, s, samp, script, select, small, span, strike, strong, style, sub, sup, table, tbody, td, textarea, tfoot, th, thead, title, tr, tt, u, ul, var



## 3. ETIQUETAS Y ATRIBUTOS

Se trata de un número de etiquetas muy grande pero no es suficiente para crear **páginas complejas**. Algunos elementos como las imágenes y los enlaces requieren cierta información adicional para estar completamente definidos.

La etiqueta `<a>` por ejemplo se emplea para incluir un **enlace** en una página. Utilizando sólo la etiqueta `<a>` no es posible establecer la dirección a la que apunta cada enlace.

**SOLUCIÓN:** **personalizar** las etiquetas HTML mediante cierta información adicional llamada **atributos**.



## 3. ETIQUETAS Y ATRIBUTOS

Se utiliza la misma etiqueta `<a>` para todos los enlaces de la página y se utilizan los atributos para indicar la dirección a la que apunta cada enlace.

```
<html>
```

```
  <head>
```

```
    <title>Ejemplo de atributos en las etiquetas</title>
```

```
  </head>
```

```
  <body>
```

```
    <p> Los enlaces son muy fáciles de indicar:
```

```
      <a>Soy un enlace incompleto, porque no tengo dirección de  
destino</a>
```

```
      <a href="http://www.google.com">Este otro enlace apunta a la  
página de Google</a>
```

```
    </p>
```

```
  </body>
```

```
</html>
```



# 3. ETIQUETAS Y ATRIBUTOS



## 3. ETIQUETAS Y ATRIBUTOS

Aunque cada una de las etiquetas HTML define sus propios atributos, algunos de los atributos son comunes a muchas o casi todas las etiquetas. Los atributos comunes se dividen en cuatro grupos según su funcionalidad:

- **Atributos Básicos:** se pueden utilizar prácticamente en todas las etiquetas HTML.
- **Atributos para Internacionalización:** los utilizan las páginas que muestran sus contenidos en varios idiomas.
- **Atributos de Eventos:** sólo se utilizan en las páginas web dinámicas creadas con JavaScript.
- **Atributos de Foco:** relacionados principalmente con la accesibilidad de los sitios web.

## + 3.1. ATRIBUTOS BÁSICOS

Los siguiente cuatro atributos básicos se pueden aplicar prácticamente a todas las etiquetas HTML:

Atributo	Descripción
<code>id = "texto"</code>	Establece un identificador único a cada elemento dentro de una página HTML
<code>class = "texto"</code>	Establece la clase CSS que se aplica a los estilos del elemento
<code>style = "texto"</code>	Establece de forma directa los estilos CSS de un elemento
<code>title = "texto"</code>	Establece el título a un elemento (mejora la accesibilidad y los navegadores lo muestran cuando el usuario pasa el ratón por encima del elemento)

## + 3.1. ATRIBUTOS BÁSICOS

La mayoría de páginas web actuales utilizan los atributos `id` y `class` de forma masiva.

Estos atributos sólo son realmente útiles cuando se trabaja con **CSS** y con *JavaScript*.

### CARACTERÍSTICAS

- El valor de estos atributos sólo puede contener **guiones medios** (-), **guiones bajos** (\_), **letras** y/o **números**, pero no pueden empezar por números.
- Los navegadores distinguen mayúsculas de minúsculas y no se recomienda utilizar letras como **ñ** y **acentos**, ya que no es seguro que funcionen correctamente en todas las versiones de todos los navegadores.

## 3.2. ATRIBUTOS PARA INTERNACIONALIZACIÓN

Son útiles para páginas que muestran sus contenidos en **varios idiomas** y para aquellas que quieren indicar de forma explícita el idioma de sus contenidos:

Atributo	Descripción
<code>lang = "codigo de idioma"</code>	Indica el idioma del elemento mediante un código predefinido
<code>xml:lang = "codigo de idioma"</code>	Indica el idioma del elemento mediante un código predefinido
<code>dir</code>	Indica la dirección del texto (útil para los idiomas que escriben de derecha a izquierda)

## + 3.3. ATRIBUTOS DE EVENTOS

Estos atributos sólo se utilizan en las páginas web que **incluyen código JavaScript** para realizar **acciones dinámicas** sobre los elementos de la página.

Cada vez que el usuario *pulsa una tecla, mueve su ratón o pulsa cualquier botón del ratón*, se produce un **evento** dentro del navegador.

Utilizando JavaScript y los siguientes atributos, es posible responder de forma adecuada a cada evento.



## 3.3. ATRIBUTOS DE EVENTOS

Atributo	Descripción	Elementos que pueden usarlo
<code>onblur</code>	Deseleccionar el elemento	<code>&lt;button&gt;</code> , <code>&lt;input&gt;</code> , <code>&lt;label&gt;</code> , <code>&lt;select&gt;</code> , <code>&lt;textarea&gt;</code> , <code>&lt;body&gt;</code>
<code>onchange</code>	Deseleccionar un elemento que se ha modificado	<code>&lt;input&gt;</code> , <code>&lt;select&gt;</code> , <code>&lt;textarea&gt;</code>
<code>onclick</code>	Pinchar y soltar el ratón	Todos los elementos
<code>ondblclick</code>	Pinchar dos veces seguidas con el ratón	Todos los elementos
<code>onfocus</code>	Seleccionar un elemento	<code>&lt;button&gt;</code> , <code>&lt;input&gt;</code> , <code>&lt;label&gt;</code> , <code>&lt;select&gt;</code> , <code>&lt;textarea&gt;</code> , <code>&lt;body&gt;</code>
<code>onkeydown</code>	Pulsar una tecla (sin soltar)	Elementos de formulario y <code>&lt;body&gt;</code>
<code>onkeypress</code>	Pulsar una tecla	Elementos de formulario y <code>&lt;body&gt;</code>
<code>onkeyup</code>	Soltar una tecla pulsada	Elementos de formulario y <code>&lt;body&gt;</code>
<code>onload</code>	La página se ha cargado completamente	<code>&lt;body&gt;</code>

## + 3.4. ATRIBUTOS DE FOCO

Cuando el usuario selecciona un elemento en una aplicación, se dice que **"el elemento tiene el foco del programa"**.

*EJEMPLO: Un usuario pincha con su ratón sobre un cuadro de texto y comienza a escribir, ese cuadro de texto tiene el foco del programa, llamado **"focus"**. Si el usuario selecciona después otro elemento, el elemento original pierde el foco y el nuevo elemento es el que tiene el foco del programa.*

Los elementos de las páginas web pueden obtener el foco de la aplicación y HTML define algunos atributos específicos para **controlar cómo se seleccionan los elementos**.



## + 3.4. ATRIBUTOS DE FOCO

Atributo	Descripción
<code>accesskey = "letra"</code>	Establece una tecla de acceso rápido a un elemento HTML
<code>tabindex = "numero"</code>	Establece la posición del elemento en el orden de tabulación de la página. Su valor debe estar comprendido entre 0 y 32.767
<code>onfocus, onblur</code>	Controlan los eventos JavaScript que se ejecutan cuando el elemento obtiene o pierde el foco

## + 4. ELEMENTOS HTML

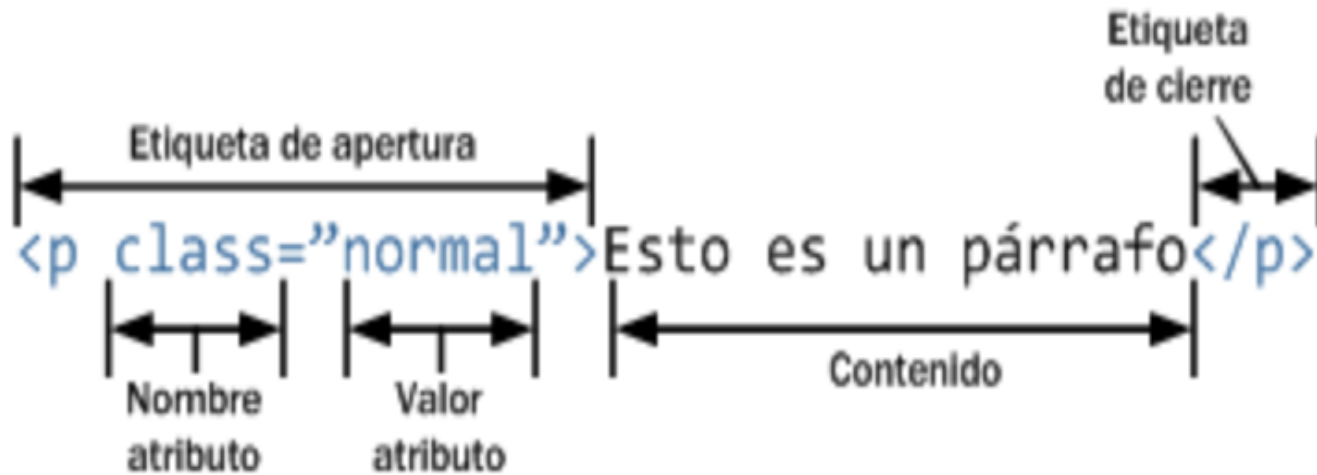
HTML define el término **elemento** para referirse a las partes que componen los documentos HTML.

Un elemento HTML es mucho más que una etiqueta, ya que está formado por:

- Una etiqueta de apertura.
- Cero o más atributos.
- Texto encerrado por la etiqueta.
- Una etiqueta de cierre.

El texto encerrado por la etiqueta es **opcional**, ya que algunas etiquetas de HTML no pueden encerrar ningún texto.

## + 4. ELEMENTOS HTML



La estructura mostrada en el esquema anterior es un elemento HTML ya que comienza con una etiqueta de apertura (`<p>`), contiene cero o más atributos (`class="normal"`), dispone de un contenido de texto (`Esto es un párrafo`) y finaliza con una etiqueta de cierre (`</p>`).

## + 4. ELEMENTOS HTML

Si una página web tiene dos párrafos de texto, la página contiene dos elementos y cuatro etiquetas (dos etiquetas `<p>` de apertura y dos etiquetas `</p>` de cierre).

El lenguaje HTML clasifica a todos los elementos en dos grupos: **elementos en línea** (*inline elements*) y **elementos de bloque** (*block elements*).

La principal diferencia es la forma en la que **ocupan el espacio disponible** en la página. Los elementos de bloque siempre empiezan en una nueva línea y **ocupan todo el espacio disponible** hasta el final de la línea, aunque sus contenidos no lleguen hasta el final de la línea. Por su parte, los elementos en línea sólo **ocupan el espacio necesario** para mostrar sus contenidos.



## 4. ELEMENTOS HTML

```
<html>
```

```
<head>
```

```
<title>Ejemplo de elementos en línea y elementos de bloque</title>
```

```
</head>
```

```
<body>
```

```
<p>Los párrafos son elementos de bloque.</p>
```

```
<a href="http://www.google.com">Los enlaces son elementos en línea</a>
```

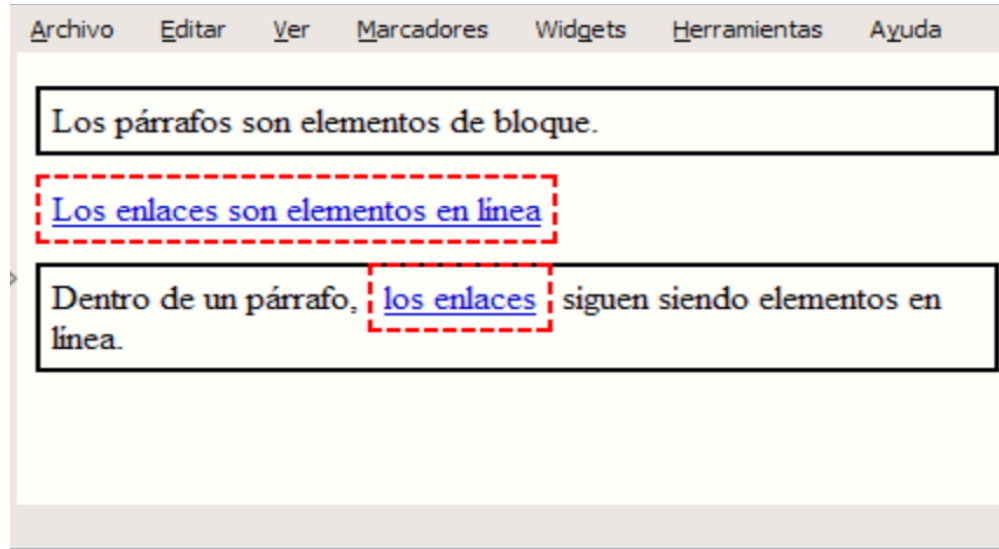
```
<p>Dentro de un párrafo, <a href="http://www.google.com">los  
enlaces</a>
```

```
siguen siendo elementos en línea.</p>
```

```
</body>
```

```
</html>
```

## + 4. ELEMENTOS HTML



- El primer párrafo contiene un texto corto. No obstante, el espacio reservado por el navegador para el primer párrafo llega hasta el final de esa línea, por lo que resulta evidente que los elementos `<p>` son elementos de bloque.
- El primer enlace tiene un texto corto. El navegador sólo reserva para el enlace el sitio necesario para mostrar sus contenidos. Si se añade otro enlace en esa misma línea, se mostraría a continuación del primer enlace. Por tanto, los elementos `<a>` son elementos en línea.
- El segundo párrafo sigue ocupando todo el espacio disponible hasta el final de cada línea y el enlace que se encuentra dentro del párrafo sólo ocupa el sitio necesario para mostrar sus contenidos.

## + 4. ELEMENTOS HTML

- **Los elementos en línea** definidos por HTML son: `a`, `abbr`, `acronym`, `b`, `basefont`, `bdo`, `big`, `br`, `cite`, `code`, `dfn`, `em`, `font`, `i`, `img`, `input`, `kbd`, `label`, `q`, `s`, `samp`, `select`, `small`, `span`, `strike`, `strong`, `sub`, `sup`, `textarea`, `tt`, `u`, `var`.
- **Los elementos de bloque** definidos por HTML son: `address`, `blockquote`, `center`, `dd`, `dt`, `dir`, `div`, `dl`, `fieldset`, `form`, `frame-set`, `h1`, `h2`, `h3`, `h4`, `h5`, `h6`, `hr`, `isindex`, `li`, `menu`, `noframes`, `nos-crypt`, `ol`, `p`, `pre`, `table`, `tbody`, `tfoot`, `th`, `thead`, `tr`, `ul`.
- Los siguientes elementos pueden ser **en línea y de bloque** según las circunstancias: `button`, `del`, `iframe`, `ins`, `map`, `object`, `script`.



## 5. SINTAXIS DE LAS ETIQUETAS XHTML

El lenguaje HTML original era muy **permisivo en su sintaxis**, por lo que era posible escribir sus etiquetas y atributos de muchas formas diferentes.

- Las etiquetas podían escribirse en mayúsculas, en minúsculas e incluso combinando mayúsculas y minúsculas.
- El valor de los atributos de las etiquetas se podían indicar con y sin comillas.
- El orden en el que se abrían y cerraban las etiquetas no era importante.

La flexibilidad de HTML puede parecer un aspecto positivo, pero el resultado final son **páginas con un código HTML desordenado**, difícil de mantener y muy poco profesional. Afortunadamente, las versiones más modernas de HTML solucionan estos problemas añadiendo ciertas normas en la forma de escribir las etiquetas y atributos.





## 5. SINTAXIS DE LAS ETIQUETAS XHTML

**1) Las etiquetas se tienen que cerrar de acuerdo a como se abren:**

*Ejemplo correcto en XHTML:*

```
<p>Este es un párrafo con <a>un enlace</a></p>
```

*Ejemplo incorrecto:*

```
<p>Este es un párrafo con <a>un enlace</p></a>
```



## 5. SINTAXIS DE LAS ETIQUETAS XHTML

**2) Los nombres de las etiquetas y atributos siempre se escriben en minúsculas:**

*Ejemplo correcto en XHTML:*

```
<p>Este es un párrafo con <a href="http://  
www.google.com">un enlace</a></p>
```

*Ejemplo incorrecto:*

```
<P>Este es un párrafo con <A HREF="http://  
www.google.com">un enlace</A></P>
```



## 5. SINTAXIS DE LAS ETIQUETAS XHTML

**3) El valor de los atributos siempre se encierra con comillas:**

*Ejemplo correcto en XHTML:*

```
<p>Este es un párrafo con <a href="http://  
www.google.com">un enlace</a></p>
```

*Ejemplo incorrecto:*

```
<p>Este es un párrafo con <a href=http://  
www.google.com>un enlace</a></p>
```

## + 5. SINTAXIS DE LAS ETIQUETAS XHTML

### 4) Los atributos no se pueden comprimir:

*Ejemplo correcto en XHTML:*

```
<dl compact="compact">...</dl>
```

*Ejemplo incorrecto:*

```
<dl compact>...</dl>
```

Este tipo de atributos en los que el nombre coincide con su valor no son muy habituales.

## + 5. SINTAXIS DE LAS ETIQUETAS XHTML

### 5) Todas las etiquetas deben cerrarse siempre:

Algunas etiquetas especiales llamadas "etiquetas vacías" no necesitan encerrar ningún texto.

La etiqueta `<br>`, se utiliza para indicar el comienzo de una nueva línea. Por sus características, la etiqueta `<br>` nunca encierra ningún contenido de texto. XHTML permite en estos casos escribir de forma abreviada una etiqueta que se abre y se cierra de forma consecutiva.

En lugar de abrir y cerrar de forma consecutiva la etiqueta (`<br></br>`) se puede utilizar **la sintaxis** `<br/>` para indicar que es una etiqueta vacía que se abre y se cierra en ese mismo punto.