

# *APLICACIONES WEB*

## **TEMA 2:** *LENGUAJES DE PRESENTACIÓN Y ESTILO*



### **Objetivos**

- ✓ Introducir el lenguaje de presentación de contenidos HTML
- ✓ Mostrar las diferencias básicas del XHTML respecto al HTML convencional
- ✓ Introducir las hojas de estilo en cascada CSS

# Índice: Tema 2.1

## **2.1 El idioma de Internet**

2.1.1 El lenguaje HTML

2.1.2 El uso de marcas HTML

2.1.3 Componentes de los documentos HTML

# Índice: Tema 2.1

## **2.1 El idioma de Internet**

### **2.1.1 El lenguaje HTML**

### 2.1.2 El uso de marcas HTML

### 2.1.3 Componentes de los documentos HTML



## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ ¿Qué es una web?

- ✓ Una página WEB no es un único objeto, está compuesto por múltiples ficheros (imágenes, video, películas flash, JavaScript ).
- ✓ Existe un fichero principal cuyo contenido es HTML y que describe el contenido de la página, tanto texto como otros elementos.
- ✓ Las WWWs es una combinación de 4 ideas básicas:
  - Hipertexto (enlaces a otros recursos).
  - Identificadores de Recursos (URLs).
  - Modelo cliente-servidor.
  - Un lenguaje de marcado (HTML).





## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ ¿Qué es HTML?

- ✓ **HTML** es una aplicación SGML (Standard Generalized Markup Language).
- ✓ **SGML** es un metalenguaje creado para definir lenguajes de marcado específicos:
  - Se usan para marcar la información con diferentes propósitos: señalar cómo debe presentarse la información, indicar el significado de cada dato, etc...
  - Para ello, los elementos de SGML se definen mediante **marcas de apertura y cierre**:  
`<nombre> . . . </nombre>`
- ✓ HTML ha ido evolucionado en sucesivas versiones desde que Tim Berners-Lee propusiera la primera versión en 1991, estableciéndose estándares para cada una.
- ✓ La versión estandarizada actual es HTML4.01 (HTML5 es aún experimental).
- ✓ El organismo internacional encargado de la estandarización de las versiones de HTML es el *World Wide Web Consortium* (**W3C**).



## 2.1 EL IDIOMA DE INTERNET

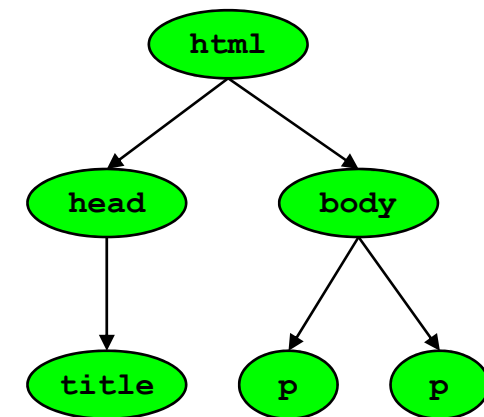
### 2.1.1 EL LENGUAJE HTML



#### ➤ Estructura básica de un documento HTML

- ✓ Todo documento HTML consta de tres secciones:
  - Declaración de la **DTD** utilizada.
  - **Cabecera** con información general sobre el documento (entre las marcas `<head>`).
  - **Cuerpo** con la descripción del contenido del documento (entre las marcas `<body>`).
- ✓ Las tres secciones se han de encontrar en ese orden.
- ✓ La cabecera y el cuerpo se encuentran dentro del elemento HTML, pero la declaración de versión no.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN"
"http://www.w3.org/TR/REC-html40/strict.dtd">
<html>
  <head> ... </head>
  <body> ... </body>
</html>
```





## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Definición de Tipo de Documento (1)

- ✓ Como todas las aplicaciones SGML, HTML está definido con una Definición de Tipo de Documento (**DTD**).
- ✓ Una DTD es un documento que recoge el conjunto de normas y restricciones que deben cumplir los documentos de un determinado tipo.
- ✓ La DTD para HTML establece la forma de construir marcas, las marcas contempladas y sus atributos, las secuencias correctas, etc....
- ✓ Para que una página HTML sea correcta y válida es imprescindible que incluya el correspondiente **DOCTYPE** que indica el DTD utilizado.
- ✓ Para HTML 4 están definidas tres tipos de DTD: estricta, transicional y para documentos con marcos.



## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Definición de Tipo de Documento (2)

- ✓ **DTD estricta:** Incluye las restricciones más severas.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN"  
"http://www.w3.org/TR/REC-html40/strict.dtd">
```

- ✓ **DTD transicional:** Permite el uso de algunos atributos relacionados con el aspecto.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN"  
"http://www.w3.org/TR/REC-html40/loose.dtd">
```

- ✓ **DTD para páginas de marcos:** Permite el uso de marcos (obsoleto).

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Frameset//EN"  
"http://www.w3.org/TR/REC-html40/frameset.dtd">
```





## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Definición de Tipo de Documento (3)

- ✓ La **validación** es el proceso que asegura que un documento escrito en un determinado lenguaje cumple con las normas y restricciones especificadas en su DTD.
- ✓ Aunque la validación no es obligatoria y las páginas web se pueden ver bien sin que sean válidas, es un proceso que permite detectar errores.
- ✓ Existen diferentes formas de validar una página creada de acuerdo a una determinada DTD:
  - Algunos editores (como Dreamweaver) tienen el validador integrado en la propia herramienta.
  - Algunos navegadores (como Firefox) permite instalar un validador como complemento.
  - El organismo W3C posee una herramienta que se puede utilizar gratuitamente a través de Internet desde: <http://validator.w3.org/>



## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Cabecera del documento (1)

- ✓ La **cabecera** es el elemento de segundo nivel delimitado por las marcas **<head>**.
- ✓ Contiene información general acerca del propio documento que no se considera como parte del contenido del documento:

**<head>**

*<!-- Zona de etiquetas META -->*

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

*<!-- Zona de título -->*

`<title>El título del documento</title>`

*<!-- Zona de recursos enlazados (CSS, JavaScript, ...) -->*

`<link rel="stylesheet" href="#" type="text/css" media="screen" />`

`<style> p {color: blue;} </style>`

`<script src="#" type="text/javascript"></script>`

**</head>**



## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Cabecera del documento (2)

- ✓ El elemento principal de la cabecera es **<title>**, que especifica el **título del documento**:
  - Este elemento es obligatorio.
  - Su contenido sólo puede ser texto.
  - El título suele aparecer en la barra superior de los navegadores.
- ✓ Además puede contener otros elementos como **metadatos**.
- ✓ La **zona de recursos enlazados** puede contener información relativa a la utilización de:
  - Estilos para dar formato al contenido de la página.
  - Lenguajes de script para dotar de dinamismo a la página.



## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Cabecera del documento (3)

- ✓ Los **metadatos** permiten incluir cualquier información relevante sobre la propia página mediante la etiqueta **<meta>**. Los metadatos más utilizados son:

- Autor del documento:

```
<meta name="author" content="Juan Pérez" />
```

- Codificación de caracteres del documento:

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

- Copyright del documento:

```
<meta name="copyright" content="librosweb.es" />
```

- Comportamiento de los buscadores:

```
<meta name="robots" content="index, follow" />
```

- Palabras clave que definen el contenido del documento:

```
<meta name="keywords" content="diseño, css, web, html" />
```

- Breve descripción del sitio:

```
<meta name="description" content="Artículos sobre diseño web" />
```



## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Cuerpo del documento

- ✓ El **cuerpo** es el elemento de segundo nivel, delimitado por las marcas `<body>`, que describe el contenido de la página, es decir, lo que ha de mostrar el navegador, y cómo lo ha de mostrar.
- ✓ En el cuerpo del documento puede haber infinidad de elementos de muy diversos tipos, anidados unos dentro de otros, y que son los que determinan el formato del documento. Se dividen en dos tipos generales:
  - A nivel de texto (**inline**), que pueden contener texto propiamente dicho y otros elementos a nivel de texto. Son elementos que no definen estructuras y se utilizan para establecer opciones de formato, crear hipervínculos o insertar objetos.
  - A nivel de bloque (**block-level**), que pueden contener elementos a nivel de texto y otros elementos a nivel de bloque. Se utilizan para estructurar el contenido, organizando la información en entidades tales como párrafos, epígrafes, tablas, listas o formularios.



## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Resumiendo (1)

- ✓ HTML es un **Lenguaje** de **Marcas** de **HiperTexto**.
- ✓ **Lenguaje** de computadora de sintaxis muy sencilla.
  - NO es un lenguaje de programación.
  - Documentos en archivos de texto simple.
  - Texto sin formato.
- ✓ Diferenciación de los distintos elementos por **marcas**.
  - Pares de marcas: **<nombre>** . . . **</nombre>**
- ✓ **Hipertexto**: vínculos con otros recursos.
  - Lugares del documento.
  - Otras páginas locales.
  - Otros sitios de Internet.



## 2.1 EL IDIOMA DE INTERNET

### 2.1.1 EL LENGUAJE HTML



#### ➤ Resumiendo (2)

- ✓ Un archivo HTML describe:
  - El contenido de una página web.
  - La forma en que se debe visualizar ese contenido en un navegador de Internet.
- ✓ Contenido de la página web:
  - **Texto**: Se encuentra directamente en el archivo HTML.
  - **Elementos multimedia**: Se encuentran en archivos aparte y se incluyen en la página web mediante instrucciones HTML.
- ✓ Extensión de los archivos web: **html**

# Índice: Tema 2.1

## **2.1 El idioma de Internet**

2.1.1 El lenguaje HTML

**2.1.2 El uso de marcas HTML**

2.1.3 Componentes de los documentos HTML





### ➤ Una primera página web

- ✓ Comencemos creando una página web con el siguiente fragmento de texto:

Ante la ley hay un guardián. Un campesino se presenta ante el guardián y pide entrar en la ley. Pero el guardián le dice que por el momento no puede dejarlo entrar. Tras reflexionar, pregunta si lo dejarán entrar más tarde.

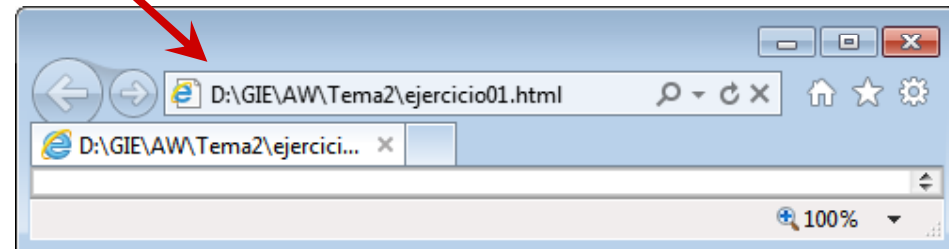
Es posible -dice el guardián-, pero ahora no.

[...]

- ✓ Para ello, creamos el archivo:
  - Abriendo el Bloc de notas.
  - Escribiendo el texto.
  - Guardando el archivo con el nombre *ejercicio01.html*

#### ➤ Visualización de la página web (1)

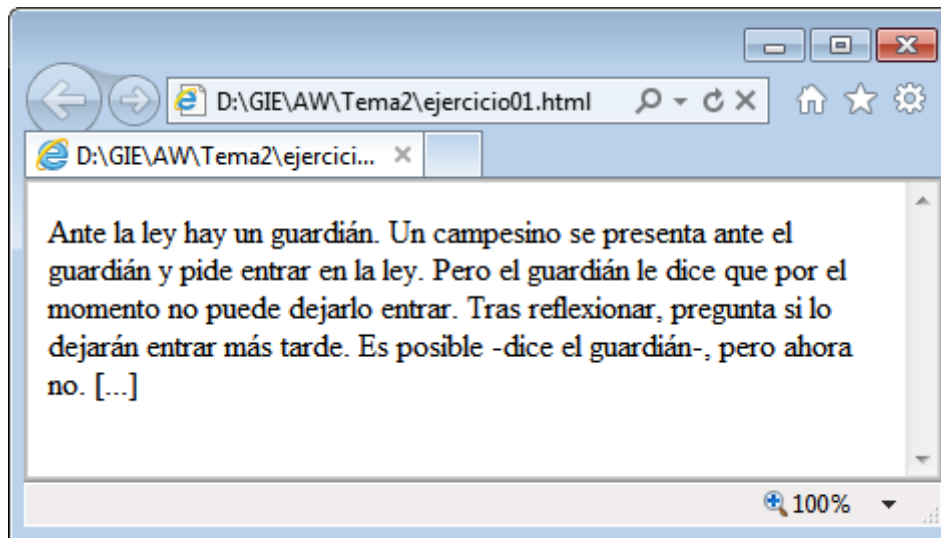
- ✓ Todos los navegadores entienden HTML, aunque la visualización puede resultar ligeramente distinta.
- ✓ Ajustándonos al estándar, no ha de haber problemas.
- ✓ Los navegadores proporcionan extensiones.
- ✓ Para cargar el archivo HTML en el navegador, podemos:
  - Escribir la ruta en el campo de dirección



- Abrir el archivo en el navegador.
- Haciendo doble click sobre el archivo en el explorador.

### ➤ Visualización de la página web (2)

- ✓ El navegador muestra el texto del archivo HTML, pero los resultados no son exactamente los deseados.



- ✓ Sin marcas, el navegador interpreta el texto como parte de un único párrafo, sin que sirva de nada partirlo en varios párrafos en el editor.
- ✓ El original tiene 3 párrafos, mientras que el visualizado tan solo consta de 1.



## 2.1 EL IDIOMA DE INTERNET

### 2.1.2 EL USO DE MARCAS HTML

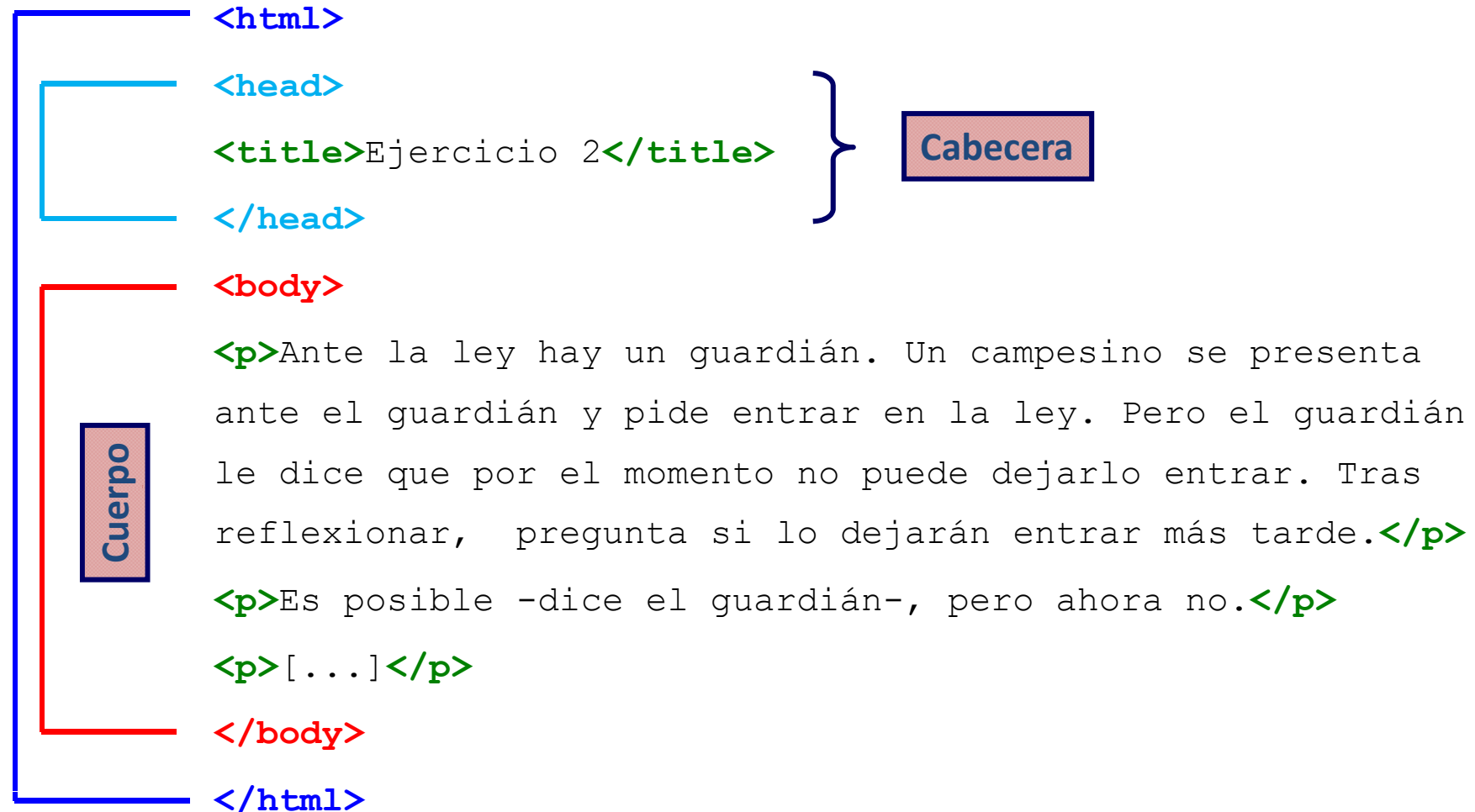


#### ➤ Las marcas en el lenguaje HTML

- ✓ ¿Para qué sirven las marcas del lenguaje HTML?
  - Para delimitar párrafos, por ejemplo. Pero también para indicar el formato del texto, para indicar el formato de la propia página, para asignar un título a la página, para establecer hipervínculos, para forzar saltos de línea, para aplicar estilos predeterminados, para incluir archivos multimedia, etc...
- ✓ ¿Qué son las marcas que define el lenguaje HTML?
  - Texto diferente al normal (el texto a visualizar), y que proporciona al navegador instrucciones sobre cómo construir la página web.
  - Cada instrucción se reconoce por el identificador (etiqueta) que tiene asignado, y puede llevar información adicional asociada en forma de atributos (opciones de la instrucción).
  - La mayoría van en parejas (apertura y cierre), encerradas entre ángulos < ... > y no se distingue entre mayúsculas o minúsculas.



#### ➤ Ejemplo de un documento con marcas (1)





## 2.1 EL IDIOMA DE INTERNET

### 2.1.2 EL USO DE MARCAS HTML

### EJERCICIO 2

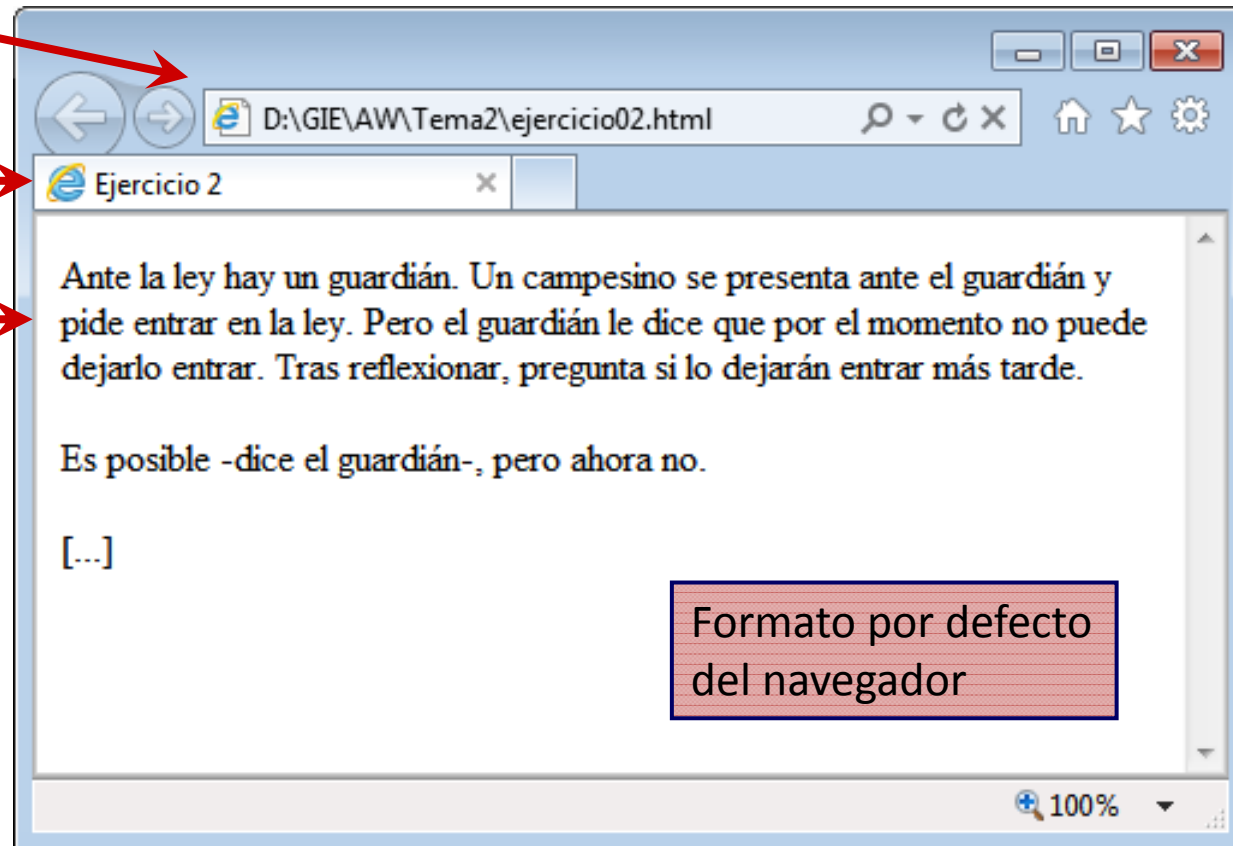


#### ➤ Ejemplo de un documento con marcas (2)

Documento

Título

Cuerpo





## 2.1 EL IDIOMA DE INTERNET

### 2.1.2 EL USO DE MARCAS HTML



#### ➤ Las marcas de un documento HTML básico

✓ `<html> ... </html>`

- Delimita la sección de descripción de la página web.
- De momento engloba todo el documento HTML.

✓ `<head> ... </head>`

- Delimita la sección de cabecera del documento.

✓ `<body> ... </body>`

- Delimita la sección de cuerpo del documento.

✓ `<title> ... </title>`

- Define el título del documento (título de la página web).
- Aparecerá en la barra de título del navegador.

✓ `<p> ... </p>`

- Engloba el texto de un párrafo.



## 2.1 EL IDIOMA DE INTERNET

### 2.1.2 EL USO DE MARCAS HTML

#### EJERCICIO 3



#### ➤ Hipertexto: La inclusión de (hiper)vínculos (1)

- ✓ Referencia que se establece encerrando el texto del vínculo entre marcas `<a> . . . </a>` y definiendo la dirección de destino con un atributo `href` dentro de la marca de apertura.

- ✓ Supongamos que añadimos el siguiente párrafo a la página del Ejercicio 2, en el que queremos que el texto **Franz Kafka** tenga un vínculo a la página destino, [http://es.wikipedia.org/wiki/Franz\\_Kafka](http://es.wikipedia.org/wiki/Franz_Kafka), a la que se salte al ser pulsado:

`<p>Franz Kafka - Traducción de J.R. Wilcok</p>`

El resultado sería:

`<p>`

`<a href="http://es.wikipedia.org/wiki/Franz_Kafka">Franz  
Kafka</a> - Traducción de J.R. Wilcok`

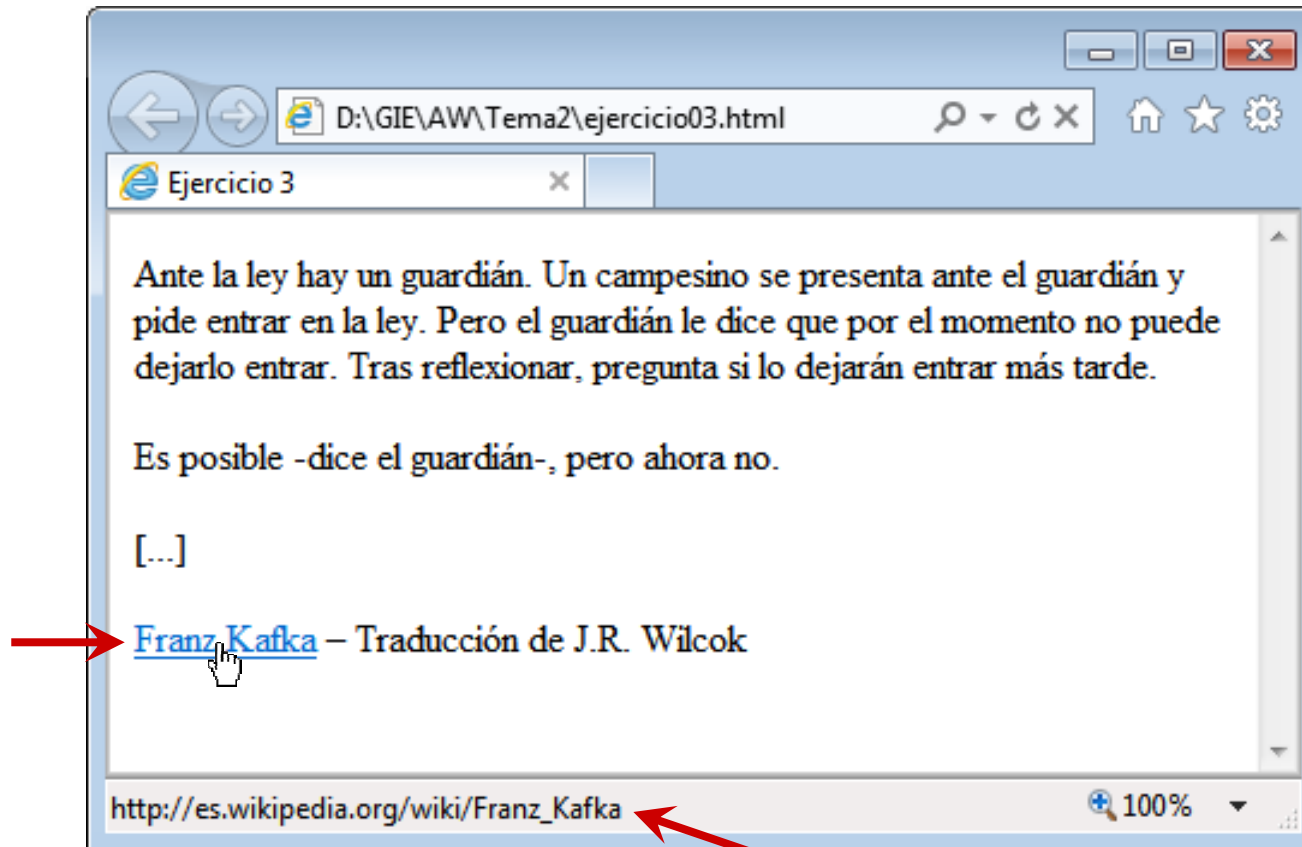
`</p>`





#### ➤ Hipertexto: La inclusión de (hiper)vínculos (2)

Hipervínculo



Destino

# Índice: Tema 2.1

## **2.1 El idioma de Internet**

2.1.1 El lenguaje HTML

2.1.2 El uso de marcas HTML

**2.1.3 Componentes de los documentos HTML**



## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML

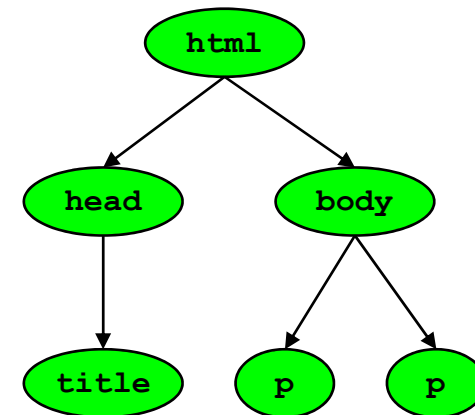


#### ➤ Estructura de los documentos HTML:

- ✓ Conjunto de elementos organizados jerárquicamente, en donde cada uno de ellos es de un tipo concreto (etiqueta):

**<etiqueta>**Contenido del elemento**</etiqueta>**

- ✓ Tipos: **html**, **head**, **body**, **title**, **p**, **a** y muchos más.
- ✓ Estructura jerárquica: unos elementos se encuentran dentro de otros.
- ✓ Existen cuatro tipos de componentes:
  - Elementos.
  - Atributos (propiedades).
  - Caracteres especiales.
  - Comentarios.





## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML



#### ➤ Elementos (1)

- ✓ Se identifican delimitándolos con un par de marcas, cuya etiqueta identifica el tipo de elemento. Entre las marcas se encuentra el contenido, que puede ser una secuencia de texto normal, referencias de caracteres y otros elementos.
- ✓ Así, puede haber anidamiento de elementos según una estructura jerárquica.
  - Existen relaciones de herencia al ser unos elementos padres de otros. Los hijos adquieren automáticamente las propiedades del padre.
  - Elemento raíz: el que está en lo más alto de la jerarquía (primer nivel o nivel superior, siempre de tipo HTML), y contiene al resto de los elementos del documento.
- ✓ A veces se puede omitir la marca de cierre aunque no es una buena práctica (menor legibilidad del código).
- ✓ Algunos elementos no tienen contenido, son elementos vacíos y no hay marca de cierre). Ejemplo: **<br>** (salto de línea)



## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML



#### ➤ Elementos (2)

##### ✓ Tipos de etiquetas HTML (1):

- Etiquetas que definen la estructura del documento: `<html>`, `<head>`, `<body>`
- Etiquetas que pueden ir en la cabecera: `<title>`, `<base>`, `<meta>`, `<style>`, `<link>`
- Etiquetas que definen bloques de texto: `<address>`, `<blockquote>`, `<div>`, `<h1>...`  
`<h6>`, `<p>`, `<pre>`, `<xmp>`
- Etiquetas de listas: `<dir>`, `<dl>`, `<dt>`, `<dd>`, `<menu>`, `<ol>`, `<ul>`, `<li>`
- Etiquetas de características del texto: `<b>`, `<basefont>`, `<big>`, `<blink>`, `<cite>`,  
`<code>`, `<em>`, `<font>`, `<i>`, `<kbd>`, `<plaintext>`, `<small>`, `<s>`, `<strike>`,  
`<strong>`, `<sub>`, `<sup>`, `<tt>`, `<u>` , `<var>`
- Etiquetas de anclas y enlaces: `<a>`
- Etiquetas de imágenes y mapas de imágenes: `<img>`, `<area>`, `<map>`



## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML



#### ➤ Elementos (3)

##### ✓ Tipos de etiquetas HTML (2):

- Etiquetas de tablas: `<table>`, `<caption>`, `<tr>`, `<td>` y `<th>`
- Etiquetas de formularios: `<form>`, `<input>`, `<select>`, `<option>`, `<textarea>`, `<keygen>`, `<isindex>`
- Etiquetas de marcos: `<frame>`, `<frameset>`, `<noframes>`
- Etiquetas de situación de contenidos: `<layer>`, `<ilayer>` , `<nolayer>`
- Etiquetas de script: `<script>`, `<noscript>` , `<server>`
- Etiquetas de applets y plug-ins: `<applet>`, `<param>`, `<embed>`, `<noembed>` , `<object>`
- Etiquetas de ajuste del texto: `<br>`, `<center>`, `<hr>`, `<multicol>`, `<nobr>`, `<spacer>`, `<span>`, `<wbr>`



## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML



#### ➤ Atributos (1)

- ✓ Los **<elementos>** (o etiquetas) pueden tener definidas propiedades.
  - Cada atributo de un elemento representa una característica del mismo que informa al navegador sobre un aspecto que ha de tener en cuenta a la hora de interpretarlo (color o fuente para el texto, destino de un hipervínculo, etc...)
- ✓ A cada **atributo** se le puede asignar un valor dentro de un conjunto de valores válidos que se le pueden asignar al mismo.
- ✓ El **"valor"** va encerrado entre comillas simples o dobles. Ejemplo:  
  
`<a href="http://huizen.dds.nl/~nfkk/">`
- ✓ Cada tipo de elemento tiene definido su propio conjunto de atributos, pero algunos tienen definido ninguno (conjunto vacío).
- ✓ Para muchas propiedades hay valores predeterminados.



## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML



#### ➤ Atributos (2)

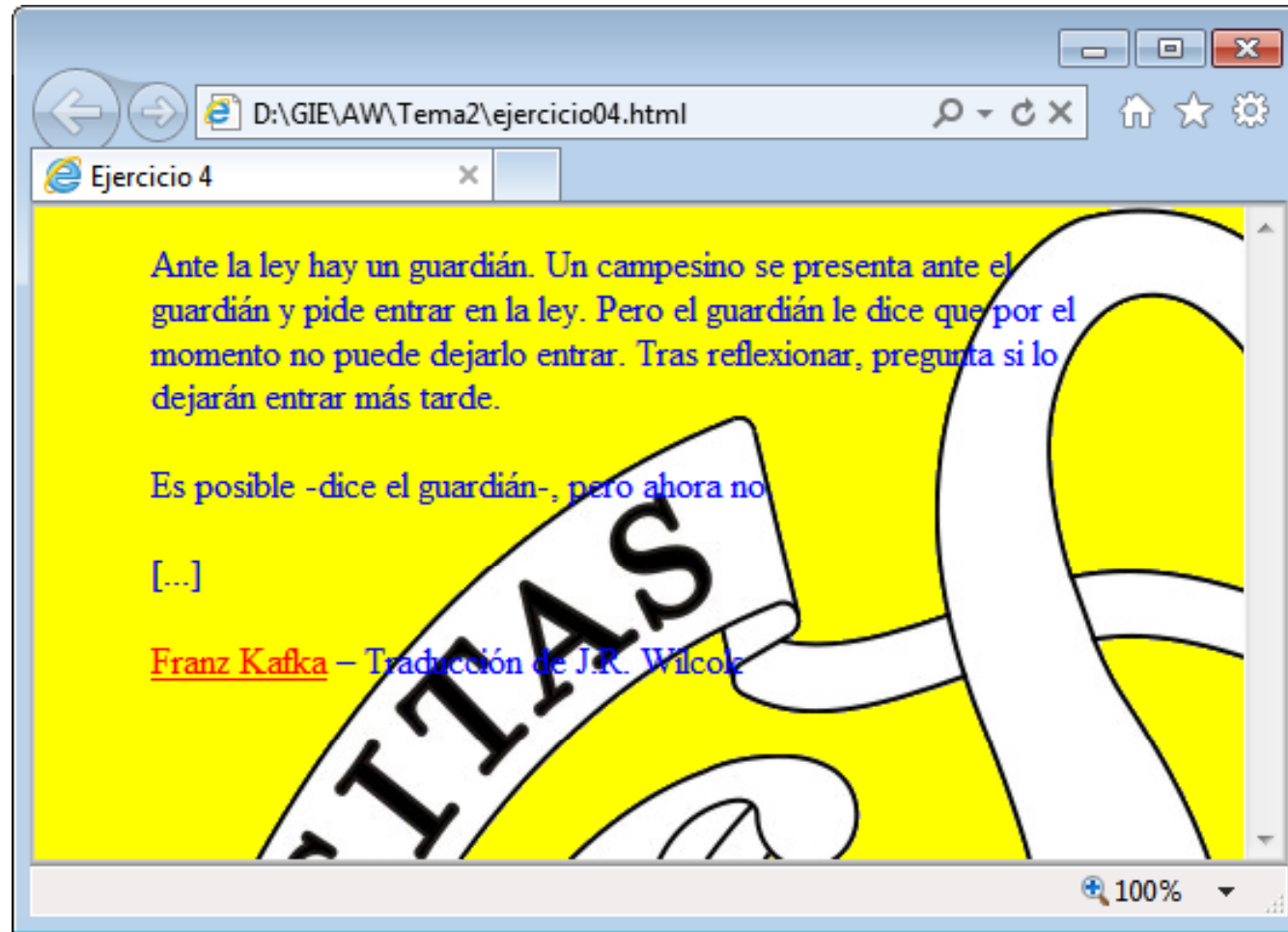
- ✓ Por ejemplo, mediante los atributos del elemento **<body>** se puede dar formato a la página en general:

- **background**: Imagen de fondo.
- **bgcolor**: Color de fondo.
- **link**: Color de los vínculos no activados.
- **vlink**: Color de los vínculos ya activados.
- **alink**: Color de los vínculos cuando son activados.
- **text**: Color del texto.
- **leftmargin**: Margen a la izquierda del texto.

```
<body background="images/imagen_ucm.gif" bgcolor="#FFFF00"  
link="#FF0000" vlink="#008080" alink="#008000" text="#0000FF"  
leftmargin="50">
```



#### ➤ Atributos del elemento <body>





## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML



#### ➤ Caracteres especiales (1)

- ✓ Algunos sistemas informáticos trabajan con 7 bits en vez de 8.
- ✓ En esos casos, si se desea trabajar con caracteres ASCII de la parte superior de la tabla (128-255), en HTML hay dos formas de codificarlos para evitar problemas:
  - Las **referencias decimales** (también llamadas secuencias de escape), que usan el formato **&#nnn;** donde **nnn** es el código ASCII en decimal del carácter. Por ejemplo, **&#193;** representa el carácter **Á**.
  - Las **referencias a entidades**, que usan el formato **&aaaa;** donde **aaaa** es una cadena de texto que representa el carácter. Por ejemplo, **&Aacute;** representa el carácter **Á** o **&ntilde;** es el carácter **ñ**.
- ✓ Además existen caracteres en la parte inferior de la tabla (0-127) que por su significado especial (**"**, **<**, **>**, **&**) también es necesario codificar. Por ejemplo, **&#60;** ó **&lt;** representan el carácter **<**.



## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML



#### ➤ Caracteres especiales (2)

Entidad	Carácter	Descripción oficial
&ntilde;	ñ	latin letter n with tilde
&Ntilde;	Ñ	latin capital n letter with tilde
&aacute;	á	a acute
&eacute;	é	e acute
&iacute;	í	i acute
&oacute;	ó	o acute
&uacute;	ú	u acute
&Aacute;	Á	A acute
&Eacute;	É	E acute
&Iacute;	Í	I acute
&Oacute;	Ó	O acute
&Uacute;	Ú	U acute
&euro;	€	euro

Entidad	Carácter	Descripción	Traducción
&lt;	<	less than	signo de menor que
&gt;	>	more than	signo de mayor que
&amp;	&	ampersand	ampersand
&quot;	"	quotation mark	comillas
&nbsp;	(espacio en blanco)	non-breaking space	espacio en blanco
&apos;	'	apostrophe	apóstrofo



## 2.1 EL IDIOMA DE INTERNET

### 2.1.3 COMPONENTES DE LOS DOCUMENTOS HTML



#### ➤ Caracteres especiales (3)

- ✓ También existen otros caracteres que son difíciles de conseguir desde el teclado, ya que no están representados (§, ¶, ©, ®). Por ejemplo, `&#169;` ó `&copy;` representa el carácter ©.

#### ➤ En HTML también es posible incluir **comentarios**:

- ✓ Facilitan la legibilidad y comprensión del código.
- ✓ El navegador los ignora.
- ✓ Van entre las marcas `<!--` y `-->`  
  
`<!-- Esto es un comentario -->`
- ✓ Pueden ir en cualquier parte excepto dentro de una marca.