Lenguajes de Marcas y sistemas de Gestión de la Información

Tema 2

HTML



1 Estándares W3C/SOAP

W3C (World Wide Web consortium)

- Su función principal es tutelar el crecimiento y organización de la web
 - El W3C se creó en 1994 por Tim BernersLee en el MIT, actual sede central del consorcio.
 - Abril de 1995, el INRIA en Francia, reemplazado por el ERCIM en el 2003 como el huésped europeo del consorcio
 - La Universidad de Kei (Shonan Fujisawa Campus) en Japón en septiembre de 1996 como huésped asiático

Formalizo el lenguaje HTML

- Lenguaje de marcan con el que se escriben páginas web.
- Al crecer el uso de la web crecen las presiones para ampliar HTML
- W3C decide no ampliar HTML
 - Crea unas reglas para que cualquiera pueda crear lenguajes de marcas, según sus necesidades
 - Se deben mantener las estructuras y sintaxis comunes que permitan la compatibilidad
 - Este conjunto de reglas es XML



1 Estándares W3C/SOAP

SOAP (Simple Object Access Protocol)

- Protocolo estándar que define como dos objetos en diferentes procesos pueden comunicarse por medio de intercambio de datos XML
- El protocolo deriva del creado por Dave Winer en 1998 llamado XML-RPC
- Soap es creado por Microsoft, IBM y iris
- Actualmente está supervisado por W3C.
- Es el protocolo mas utilizado en servicios WEB
- Soap es un paradigma de mensajería de una dirección sin estado
 - Puede ser usado como parte de otros protocols mas complejos y completos



2 Evolución y versiones HTML al XHTML

HTML es el lenguaje utilizado para crear la mayor parte de las páginas web

- Es un estándar reconocido en todos los navegadores
 - Todos ellos visualizan una página HTML de forma muy similar independientemente del sistema operativo sobre el que se ejecutan

Versiones:

- HTML fue un sistema de hipertexto para compartir documentos electrónicos en 1980.
- HTML 2.0 fue la primera versión oficial de HTML el IETF publicó el estándar en septiembre de 1995
- HTML 3.2 se publicó el 14 de Enero de 1997 por el W3C. Incorpora los applets de Java y texto alrededor de las imágenes
- HTML 4.0 se publicó el 24 de Abril de 1998. Entre las novedades que presenta se encuentran las hojas de estilos CSS y la posibilidad de incluir pequeños programas en las páginas web HTML 5.0. Se publicó en octubre de 2014. Especifica dos variantes de sintaxis para HTML: una «clásica», HTML (text/html), conocida como HTML
- HTML 5, y una variante XHTML conocida como sintaxis XHTML 5 que deberá servirse con sintaxis XML (application/xhtml+xml). Esta es la primera vez que HTML y XHTML se desarrollan en paralelo



- La estructura de una página HTML es coherente con un documento XML.
 - Prologo
 - Ejemplar



Prologo

- Todo documento ha de indicar al navegador el tipo de documento que es y la versión HTML utilizada
- Version HTML 4.0 hay 3 tipos:
 - HTML 4.0 Strict
 - HTML 4.0 Transitional
 - HTML 4.0 Framese



- HTML 4.0 Strict.
 - Es la DTD utilizada por defecto con HTML 4.0.
 - En estos documentos no se permite el uso de los elementos declarados deprecated en otras versiones o Recomendaciones HTML.
 - La declaración del tipo de documento correspondiente es:
 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 //EN" "http://www.w3.org/TR/REC-html40/strict.dtd">
- HTML 4.0 Transitional.
 - Permite el uso de todos los elementos que permite el HTML 4.0
 - Strict, además de los elementos deprecated.
 - La declaración del tipo de documento correspondiente es:

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

- HTML 4.0 Frameset.
 - Es una variante de HTML 4.0 Transitional para documentos que usan frames.
 - En estos documentos el elemento body hay que reemplazarlo por un elemento frameset.
 - La declaración del tipo de documento correspondiente es
 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Frameset//EN" "http://www.w3.org/TR/REC-html40/frameset.dtd">



Ejemplar

- Todo documento HTML está delimitado por las etiquetas html>
- El ejemplar se divide en 2 partes

Cabecera

- Delimitada por las etiquetas <head> y </head>.
- Contiene la información sobre el título de la página, el autor, palabras clave, etc.
- Dentro de esta sección es obligatorio definir el título del documento,
 - Para ello se usan las etiquetas <title> </title>.
 - Está información no se presentará en la ventana del navegador, salvo el título que aparecerá en la barra de título de la parte superior

Cuerpo

- Contiene la información que se va a presentar en la pantalla.
- Está limitado por las etiquetas <body> y </body>
 - Los documentos de tipo HTML 4.0 Frameset donde éstas se sustituyen por <frameset> y </frameset>



- Un documento HTML está formado por etiquetas y atributos.
- Al igual que en XML las etiquetas pueden ser de apertura, <etiqueta>, o de cierre, </etiqueta>.
- Una de las **diferencias** con XML es que la cantidad de etiquetas de HTML está **limitada** a aquellas que están definidas por el lenguaje.
 - Aunque HTML define una gran cantidad de etiquetas, estas no son suficientes para crear páginas complejas ya que la definición completa de ciertos elementos, como las imágenes y los enlaces, requiere información adicional.
 - Como no es posible crear una etiqueta por cada elemento diferente, se añade la información adicional a las etiquetas mediante los atributos dando lugar a los elementos
 - Para cada uno de los atributos hay definido un conjunto de valores que se le puede asignar
 - si el valor de un atributo no es válido, el navegador le ignora.
 - Cada una de las etiquetas HTML define los atributos que puede utilizar, aunque algunos de ellos son comunes a muchas etiquetas



- Clasificación de los atributos comunes según su funcionalidad
 - Atributos básicos

Atributo	Descripción		
name = "texto"	Permite asignar el nombre "texto" a un objeto HTML		
title = "texto"	Asigna un título a un elemento HTML, mejorando así la accesibilidad. Dicho título es mostrado por los navegadores cuando el usuario pasa el ratón por encima del elemento. Es especialmente útil con los elementos: a, link, img, object, abbr y acronym		
id = "texto"	Permite identificar al elemento HTML sobre el que se aplica de forma única mediante el identificador "texto". Sólo es útil cuando se trabaja con CSS y con Javascript. No pueden empezar por números y sólo puede contener letras, números, guiones medios y/o guiones bajos.		
style = "texto"	Permite aplicar al elemento HTML el estilo "texto" directamente.		
class = "texto"	Permite aplicar al elemento HTML el estilo "texto" definido en las CSS. No pueden empezar por números y sólo puede contener letras, números, guiones medios y/o guiones bajos.		



Atributos para internacionalización

• Las utilizan las páginas que muestran sus contenidos en varios idiomas

Atributo	Descripción					
dir	Indica la dirección del texto por lo que sólo puede tomar dos valores: Itr (left to right) de izquierda a derecha. Es el valor por defecto. rtl (right to left) de derecha a izquierda.					
	Especifica el idioma del encuentran en el documer		6, algunos de los valore	es posible	es son:	de este atributo s
lang = "codigo"		en	Inglés (Gran Bretaña)	Código	Español	
ang codigo		en-US	Inglés americano	fr	Francés	
		ja	Japones	fr-CA	Francés de Canada	
xml:lang = "codigo"	Especifica el idioma del e	elemento me	diante un código definido	o según l	a recomendación RFC	1766.



- Atributos para internacionalización
- Los utilizan las páginas que muestran sus contenidos en varios idiomas o aquellas que quieren

Descripción					
Indica la dirección del texto por lo que sólo puede tomar dos valores: Itr (left to right) de izquierda a derecha. Es el valor por defecto. rtl (right to left) de derecha a izquierda.					
	RFC 176	6, algunos de los valore	es posible	s son:	de este atributo se
	en		3	Español	
	en-US	Inglés americano	fr	Francés	
	ja	Japones	fr-CA	Francés de Canada	
	(left to right) de izquierda a (right to left) de derecha a specifica el idioma del ele	(left to right) de izquierda a derecha (right to left) de derecha a izquierda specifica el idioma del elemento m ncuentran en el documento RFC 176 Código en en-US	(left to right) de izquierda a derecha. Es el valor por defect (right to left) de derecha a izquierda. specifica el idioma del elemento mediante un código pre ncuentran en el documento RFC 1766, algunos de los valore Código Idioma en Inglés (Gran Bretaña) en-US Inglés americano	(left to right) de izquierda a derecha. Es el valor por defecto. (right to left) de derecha a izquierda. specifica el idioma del elemento mediante un código predefinido. ncuentran en el documento RFC 1766, algunos de los valores posible Código Idioma Código en Inglés (Gran Bretaña) es en-US Inglés americano fr	(left to right) de izquierda a derecha. Es el valor por defecto. (right to left) de derecha a izquierda. specifica el idioma del elemento mediante un código predefinido. Los posibles valores ncuentran en el documento RFC 1766, algunos de los valores posibles son: Código Idioma Código Idioma en Inglés (Gran Bretaña) es Español en-US Inglés americano fr Francés



- Atributos de eventos y atributos para elementos que pueden obtener foco
 - Solo usado en paginas web dinámicas
 - Pagina web dinámica es aquella que se ha creado con JavaScript



4,1 Elementos HTML

- Elemento HTML esta formado por :
 - Una etiqueta de apertura
 - Cero o más atributos
 - Texto encerrado por la etiqueta
 - · Opcional, no todas las etiquetas pueden encerrar texto
 - Etiqueta de cierre
- Según el modo en que ocupan el espacio disponible en la página, los elementos pueden ser de dos tipos
 - Elementos en línea
 - Solo ocupan el espacio necesario para mostrar sus contenidos.
 - Su contenido puede ser texto u otros elementos en línea
 - Elementos de bloque
 - · Los elementos de bloque siempre empiezan en una nueva línea
 - Ocupan todo el espacio disponible hasta el final de la línea
 - Su contenido puede ser texto, elementos en línea u otros elementos de bloque
- Existen elementos cuyo comportamiento puede ser en línea o de bloque según las circunstancias



4,1 Elementos HTML



4.1 Elementos HTML





4.1 Elementos HTML

- Tipos
 - Elementos de la estructura básica de un documento HTML
 - Elementos de la sección de cabecera
 - Elementos de formato al texto en párrafos
 - Elementos de Listas
 - Elementos de Tablas
 - Elementos de Formulario
 - Elementos Frame
 - · Otros elementos



4.1.1 Elementos de la estructura básica

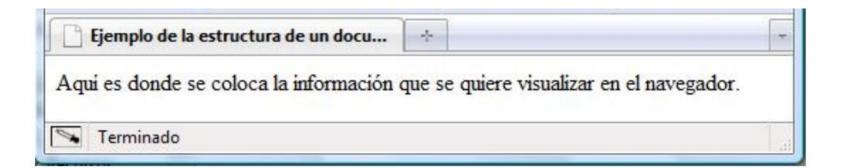
• La estructura básica de in documento viene determinada por las siguientes etiquetas:

Elemento	Descripción
html	Documento HTML.
head	Cabecera del documento.
body	Cuerpo del documento. Permite definir formatos que se aplican a los elementos de la página de manera global, como son el color del fondo del texto, los márgenes, el color de los enlaces,



4.1.1 Elementos de la estructura básica

• Ejemplo:





4.1.2 Elementos de la sección de cabecera

Elementos contenedores

Elemento	Descripción
title	Título del documento.
script	Script incrustado. Su contenido ha de ir situado entre las marcas de comentarios ya que no ha de ser interpretado.
style	Estilo aplicado al documento utilizando CSS. Su contenido ha de ir situado entre las marcas de comentarios ya que no ha de ser interpretado.

• Elementos no contenedores

Elemento	Descripción
base	URI base del documento
isindex	Prompt de entrada de datos.
link	Enlaces a documentos externos de librerías
meta	Información que agiliza la búsqueda del documento en buscadores.



4.1.3 Elementos que dan formato al texto de un párrafo

Elemento	Descripción
р	Delimita los párrafos
hi	Encabezado de nivel i, donde i es un número entero entre 1 y 6, ambos inclusive. El tamaño de la letra del encabezado es mayor cuanto menor sea el valor de i. No deben de usarse estas etiquetas para formatear texto. Sólo estarán bien usadas para designar títulos de párrafos.
b	Indica que el texto que está en ese elemento se le pondrá en negrita.
i	Indica que el texto que está en ese elemento se le pondrá en itálica ó cursiva.
u	Indica que el texto que está en ese elemento se le pondrá subrayado.
sup	Indica que el texto que está en ese elemento es un supraíndice.
sub	Indica que el texto que está en ese elemento es un subíndice.
strong	Indica que el texto que está en ese elemento estará resaltado. Habitualmente los navegadores resaltan el texto poniéndolo en negrita aunque podría haber algún navegador que resaltase el texto poniéndolo en cursiva y en naranja.

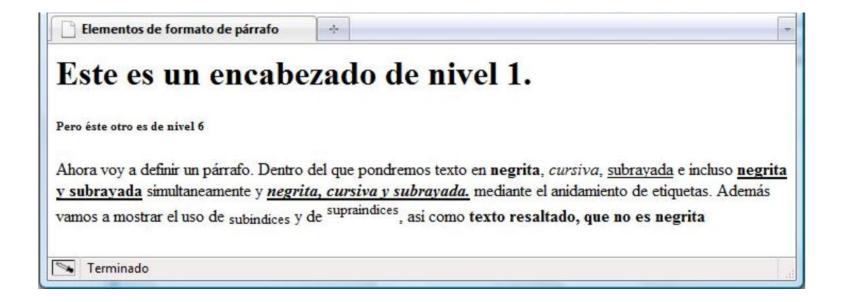


4.1.3 Elementos que dan formato al texto de un párrafo

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 transitional//EN" "http://www.w3.org/TR/REC-html40/loose.dtd">
<html>
   <head>
       <title>Elementos de formato de p&aacute; rrafo</title>
   </head>
   <body>
       <h1>Este es un encabezado de nivel 1.</h1>
       <h6>Pero &eacute; ste otro es de nivel 6</h6>
       Ahora voy a definir un pá rrafo. Dentro del que pondremos texto en <b > negrita </b >,
       <i>cursiva</i>, <u>subrayada</u> e incluso <b><u>negrita y subrayada</u></b> simultaneamente y
       <b><i>va>negrita, cursiva y subrayada.</u></b> mediante el anidamiento de etiquetas.
       Además vamos a mostrar el uso de <sub>sub&iacute;ndices</sub> y de <sup>supra&iacute;ndices</sup>,
        así como <strong>texto resaltado, que no es negrita</strong>
       </body>
</html>
```



4.1.3 Elementos que dan formato al texto de un párrafo





4.1.4 Elementos de Listas

- Tipos de listas:
 - Ordenada
 - Desordenada
 - Listas de definici

Elemento	Descripción
ul	Delimita los elementos que forman una lista desordenada
ol	Delimita los elementos que forman una lista ordenada
li	Indica cada uno de los elementos de una lista
dl	Delimita los elementos que forman una lista de definición
dt	Cada uno de los términos que se definen de una lista de definición.
dd	Cada una de las definiciones de una lista de definición.

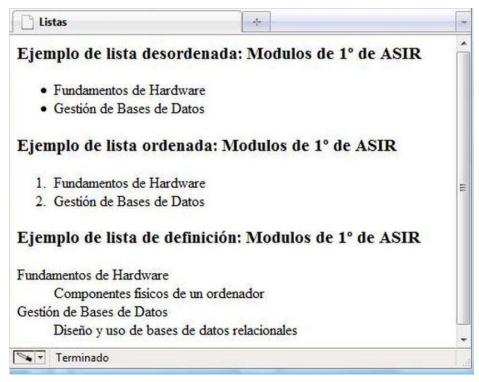


4.1.4 Elementos de Listas

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 transitional//EN" "http://www.w3.org/TR/REC-html40/loose.dtd">
<html>
   <head>
       <title>Listas</title>
   </head>
       <h3>Ejemplo de lista desordenada: Modulos de 1º de ASIR</h3>
       Fundamentos de Hardware
               Gestió n de Bases de Datos
       <h3>Ejemplo de lista ordenada: Modulos de 1º de ASIR</h3>
        Fundamentos de Hardware
           Gesti&oacute: n de Bases de Datos
       <h3>Ejemplo de lista de definici&oacute;n: Modulos de 1° de ASIR</h3>
       <dl> <dt>Fundamentos de Hardware</dt>
           <dd>Componentes f&iacute; sicos de un ordenador</dd>
           <dt>Gesti&oacute; n de Bases de Datos</dt>
           <dd>Diseño y uso de bases de datos relacionales</dd>
       </dl>
   </body>
</html>
```



4.1.4 Elementos de Listas





4.1.5 Elementos de tablas

Elemento	Descripción
table	Delimita el contenido de una tabla.
tr	Delimita cada una de las líneas de la tabla.
td	Delimita el contenido de cada celda de la tabla.
colgroup	Permite agrupar columnas.
tbody	Permite agrupar líneas de la tabla.
thead	Define la línea cabecera de la tabla.
th	Delimita cada una de las celdas de la cabecera
tfoot	Define la fila pie de la tabla.



4.1.5 Elementos de tablas

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 transitional//EN" "http://www.w3.org/TR/REC-html40/loose.dtd">
<html>
  <head>
     <title>Tablas</title>
  </head>
  <body>
     <h3>Ejemplo de tabla: Notas del m&ocute; dulo de LMSGI de 1º de ASIR</h3>
    <thead>  Eval LMSGI  
       <tfoot align="center">  Media 8  
        Primera 6 
          Segunda 7 
          Tercera 8 
       </body>
</html>
```



4.1.5 Elementos de tablas

Ejempl le ASII		la: Notas del m&ocutedulo de LMSGI de 1°
Eval	LMSGI	
Primera	6	
Segunda	7	
Tercera	8	
Media	8	



4.1.6 Elementos de Formulario

Elemento	Descripción			
form	Delimita el contenido del formulario			
input	Caja de texto para texto corto. Dependiendo del valor que tome el atributo type de este elemento podemos estar ante un texto sin más, un campo de texto donde al escribir no se visualice el contenido si no que escriba asteriscos, un botón de radio que el usuario podrá elegir, una opción que el usuario podrá activar, botones,			
textarea	Caja de texto para texto largo.			
select	Crea un menú desplegable que permite elegir de una lista de opciones que contiene el elemento.			
option	Delimita cada una de las opciones de un menú desplegable que le contiene.			
button	Permite definir un botón. Su principal ventaja frente a los botones hechos con input es que este elemento permite introducir en el botón cualquier otro elemento de HTML, como por ejemplo imágenes.			
fieldset	Permite agrupar elementos de un un formulario.			
legend	Permite poner un título al fieldset.			
label	Etiqueta de un campo del formulario.			



4.1.6 Elementos de Formulario

```
kform action="/my-handling-form-page" method="post">
<l
 <1i>>
   <label for="name">Nombre:</label>
   <input type="text" id="name" name="user name">
 <1i>>
   <label for="mail">Correo electrónico:</label>
   <input type="email" id="mail" name="user mail">
 \langle li \rangle
   <label for="msq">Mensaje:</label>
   <textarea id="msg" name="user message"></textarea>
 </form>
```



4.1.6 Elementos de Formulario

- Los elementos <label>, <input> y <textarea>
 - textarea: multilínea

```
kform action="/my-handling-form-page" method="post">
 <111>
 <1i>>
   <label for="name">Nombre:</label>
   <input type="text" id="name" name="user name">
  <1i>>
   <label for="mail">Correo electrónico:</label>
   <input type="email" id="mail" name="user mail">
 <1i>)
   <label for="msq">Mensaje:</label>
   <textarea id="msg" name="user message"></textarea>
 </form>
```



4.1.6 Elementos de Formulario

• El elemento <button>

```
  <button type="submit">Envie su mensaje</button>
```



4.1.7 Elementos Frame

 Frame: define la organización de los marcos dentro de la ventana de usuario

Elemento	Descripción	
frameset	Define la partición de la ventana del navegador en marcos. Sólo puede partirse en filas o en columnas. Para partir la ventana del navegador en filas y columnas hay que anidar frames.	
Frame	Define un marco que contiene información	

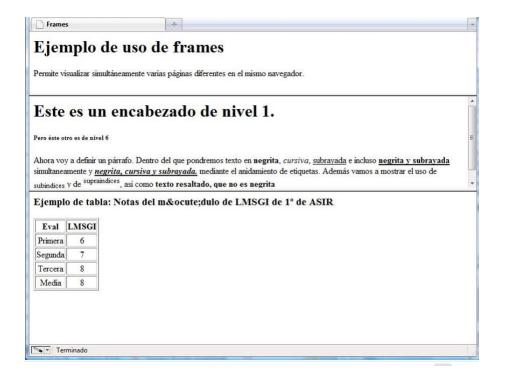


4.1.7 elementos Frame

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Frameset//EN" "http://www.w3.org/TR/REC-html40/frameset.dtd">
<html>
    <head> <title>Frames</title> </head>
    <frameset rows="20%, 30%, *">
            <frame src="LMSGI02_Ejemplo07_1.html" /> <frame src="LMSGI02_Ejemplo03.html" />
           <frame src="LMSGI02_Ejemplo05.html" />
    </frameset>
</html>
 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 transitional//EN" "http://www.w3.org/TR/REC-html40/loose.dtd">
 <html>
     <head> <title>Frames 2</title> </head>
     <body>
         <h1>Ejemplo de uso de frames</h1>
         Permite visualizar simultá neamente varias pá ginas diferentes en el mismo navegador.
     </body>
 </html>
```



4.1.7 elementos Frame





4.1.8 Otros elementos

Elemento	Descripción
a	Permite definir un enlace a una página web, un archivo o una dirección de correo.
img	Permite insertar una imagen en una página web. Es obligatorio utilizar el atributo src para determinar el path del fichero de imagen que queremos insertar.
abbr	Indica una forma abreviada.
acronym	Indica un acrónimo.
blockquote	Contiene un bloque de texto con sangría.
q	Contiene una cita por lo que el navegador le añade las marcas de citación
br	Inserta una línea en blanco. No tiene etiqueta de cierre, sólo se abre.



- XHTML es una adaptación del lenguaje HTML al XML
 - HTML tiene una sintaxis muy permisiva, lo que permite crear etiquetas y atributos de muchas formas => páginas con código desordenado
 - Un código desordenador es difícil de mantener y poco profesional
- XHTML añade normas en la forma de escribir las etiquetas



• XHTML: Diferencias sintácticas y estructurales

- Condiciones:
 - El elemento raíz del documento debe ser < html>.
 - El elemento raíz del documento debe indicar el espacio nominal XHTML usando el atributo xmlns.
 - El espacio nominal para XHTML es http://www.w3.org/1999/xhtml
 - Debe haber una declaración DOCTYPE en el prólogo del documento.
 - El identificador público incluido en la declaración DOCTYPE debe hacer referencia a alguna de las tres DTD definidas por el W3C usando el Identificador Formal Público correspondiente:



• XHTML: Diferencias sintácticas y estructurales

- El identificador público incluido en la declaración DOCTYPE debe hacer referencia a alguna de las tres DTD definidas por el W3C usando el Identificador Formal Público correspondiente:
 - <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
 - <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
 - <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">



XHTML: Restricciones básicas

- Las etiquetas se tienen que cerrar en orden inverso al que se abren, nunca pueden solaparse.
- Los nombres de las etiquetas y atributos siempre se escriben en minúsculas.
- El valor de los atributos, incluso los numéricos, siempre se encierra entrecomillas.
- Los atributos en los que el nombre coincide con su valor, no pueden darse el valor por entendido, es decir, no se pueden comprimir. Este tipo de atributos no son muy habituales.
- Todas las etiquetas deben cerrarse siempre. XHTML permite que en lugar de abrir y cerrar de forma consecutiva la etiqueta (

 la sintaxis
para indicar que es una etiqueta vacía que se abre y se cierra en ese mismo punto.



XHTML: Ventajas

- Compatibilidad parcial con navegadores antiguos: la información se visualiza, aunque sin formato.
- Un mismo documento puede adoptar diseños radicalmente distintos en diferentes apartados.
- Sencillez a la hora de editar y mantener el código.
- Es compatible con los estándares que está desarrollando el W3C como recomendación para futuros agentes de usuario o navegadores.
- Los documentos escritos conforme a XHTML 1.0 presentan mejor rendimiento en las actuales herramientas web que aquellos escritos conforme a HTML.
- La separación de los contenidos y su presentación hace que los documentos XHTML se adapten mejor a las diferentes plataformas: pantallas de ordenador, pantallas de dispositivos móviles, ...
- Como es XML se pueden utilizar fácilmente herramientas creadas para procesar documentos XML genéricos (editores, XSLT, etc.).



XHTML: Desventajas

- Algunos navegadores antiguos no son totalmente compatibles con los estándares, lo que hace que las páginas no siempre se muestren correctamente. Esto cada vez es menos problemático ya que estos navegadores van cayendo en desuso.
- Muchas herramientas de diseño web aún no generan código XHTML correcto.

