

1. Introducción

El lenguaje **XHTML** (EXtensible HyperText Markup Language) es muy similar al lenguaje **HTML**. De hecho, no es más que una adaptación de **HTML** al lenguaje **XML**, el estándar **XHTML** 1.0 sólo añade pequeñas mejoras y modificaciones menores al estándar **HTML 4.01**, por lo que este último está prácticamente incluido en el primero, lo que hace que pasar del **HTML 4.01 Strict** a **XHTML** no requiere casi ningún cambio.

El lenguaje **HTML** tiene una sintaxis muy permisiva, por lo que es posible escribir sus etiquetas y atributos de muchas formas diferentes. Las etiquetas, por ejemplo, podían escribirse en mayúsculas, en minúsculas e incluso combinando mayúsculas y minúsculas. El valor de los atributos de las etiquetas se pueden indicar con o sin comillas. Además, el orden en el que se abrían y cerraban las etiquetas no era importante.

La flexibilidad de **HTML** da lugar a páginas con un código desordenado, difícil de mantener y muy poco profesional. **XHTML** soluciona estos problemas añadiendo ciertas normas en la forma de escribir las etiquetas y atributos.

2. Diferencias sintácticas y estructurales con HTML

El esquema básico del documento, para considerarse conforme a la especificación deberá cumplir las siguientes 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:

```
<!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">
```

Restricciones básicas que introduce XHTML respecto a HTML en la sintaxis de sus etiquetas:

- 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 entre comillas.
- Los atributos en los que el nombre coincide con su valor, no puede 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 (

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

Otras restricciones:

Además de las cinco restricciones básicas, **XHTML** incluye otros cambios más avanzados respecto a **HTML**, entre ellas:

- Antes de acceder al valor de un atributo, se eliminan todos los espacios en blanco que se encuentran antes y después del valor. Además, se eliminan todos los espacios en blanco sobrantes dentro del valor de un atributo.
- El código **JavaScript** debe encerrarse entre unas etiquetas especiales (**<script>**) para evitar que el navegador interprete de forma errónea caracteres como **&** y **<**.
- Las páginas **XHTML** deben prescindir del atributo **name** en su lugar, siempre debe utilizarse el atributo **id**.
- **XHTML** es necesario separar el formato del contenido. Los párrafos deben separarse consistentemente y las cabeceras **h1-h6** sólo deben usarse para destacar los diferentes apartados. Es recomendable dar el formato a los datos por medio del uso de las hojas de estilo en cascada (CSS).

3. Ventajas e inconvenientes de XHTML sobre HTML

- **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.).
- **Inconvenientes:**
 - 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.