

¿Qué es HTML5?

Lo que hoy conocemos por HTML5 ha tenido una historia relativamente turbulenta. Como ya sabemos, HTML es el lenguaje de marcado predominante para describir el contenido o los datos en la World Wide Web. HTML5 es la última versión de este lenguaje, e incluye nuevas características, mejoras en las características existentes, y scripting basado en APIs.

Dicho esto, HTML5 no es una reformulación de las versiones anteriores del lenguaje, de hecho, incluye todos los elementos válidos tanto en HTML 4 como en XHTML 1.0. Por otra parte, ha sido diseñado con unos principios fundamentales en mente para asegurarse de que funciona en la mayoría de las plataformas existentes en el mercado, es compatible con los navegadores antiguos y se ocupa de los errores de una forma amigable.

En primer lugar, HTML5 incluye nuevos elementos de marcado cuya semántica está asociada con el significado de los contenidos que vamos a introducir en ellos. Por ejemplo, en un blog solemos tener entradas o artículos del mismo, normalmente estas entradas las introduciremos mediante la etiqueta <div>, a la cual le asociaremos una clase ¿No sería mejor introducir esos artículos mediante una etiqueta específica para ello, por ejemplo, una etiqueta <article> ? HTML5 nos proporciona nuevos elementos cuya semántica está asociada a los elementos más utilizados en el diseño web.

El término "HTML5", además, ha sido utilizado para referirse a una serie de nuevas tecnologías y APIs, como son: el dibujo con el elemento <canvas> , almacenamiento fuera de línea, el nuevo <video> y <audio> , funcionalidades de arrastrar y soltar, microdatos, etc. A lo largo del curso, veremos algunas de estas tecnologías.

¿Cómo hemos llegado hasta aquí?

La industria del diseño web se ha desarrollado en un periodo de tiempo relativamente corto. Hace doce años, un sitio web que incluía imágenes y un diseño llamativo era considerado "de primera línea" en términos de contenido web.

Ahora, el panorama es muy diferente. Son cada vez más y más comunes aplicaciones web sencillas que muestran datos en función de resultados basados en peticiones Ajax y que dependen de scripting del lado del cliente para la funcionalidad crítica. En la actualidad, cada vez encontramos más Sitios Web que son semejantes a aplicaciones de escritorio.

En la evolución del lenguaje HTML, HTML4, finalmente, dio paso a XHTML, que es en realidad HTML 4 con una sintaxis estricta de estilo XML. Hoy en día, HTML4 y XHTML son de uso general, pero HTML5 está ganando terreno.

HTML5 comenzó originalmente como dos especificaciones diferentes: Web Forms 2.0 y Aplicaciones Web 1.0. Ambas eran consecuencia de los cambios que se dieron en el paisaje web y la necesidad de que las aplicaciones web fueran fáciles de mantener y más eficientes. Finalmente, las dos especificaciones se fusionaron para formar lo que hoy conocemos como HTML5.

¿Qué es y para qué sirve XHTML?

XHTML significa lenguaje de marcado de hipertexto extendido. Se utiliza para generar documentos y contenidos de hipertexto generalmente publicados en la Web.

En general, la sintaxis XHTML es más "estricta", en el sentido de imponer restricciones en la forma de escribir etiquetas, atributos o valores, mientras que la sintaxis HTML es más "flexible".

Si un documento contiene errores de sintaxis (HTML o XHTML) se dice que es un documento inválido. Si las páginas web se sirven al navegador como application/xhtml+xml, el navegador debe rechazar las páginas inválidas, pero si se sirven al navegador como text/html, los navegadores intentan mostrar el documento aunque contenga errores. Aunque los navegadores a menudo consiguen mostrar documentos inválidos, se aconseja validar y corregir los documentos para asegurar que los navegadores puedan mostrarlos correctamente.

Generalidades

En XHTML para que un documento se muestre en el navegador tiene que ser "bien formado", es decir, tiene que cumplir todas las reglas sintácticas. En HTML, los documentos se muestran en el navegador aunque contengan errores sintácticos (la visualización puede ser correcta o incorrecta, dependiendo del tipo de errores de la página).

Veamos algunos ejemplos:

En XHTML no puede haber texto no incluido en alguna etiqueta, pero en HTML puede haberlo:

HTML ✓	<code><p>AAAA</p></code>
XHTML ✗	<code>BBBB</code> <code><p>CCCC</p></code>

En XHTML, los nombres de elementos y atributos deben escribirse en minúsculas. En HTML no es necesario.

HTML ✓	<code><p class="xxx">AAAA</p></code>
XHTML ✓	
HTML ✓	<code><P>AAAA</P></code>
XHTML ✗	
HTML ✓	<code><p CLASS="xxx">AAAA</p></code>
XHTML ✗	

Los documentos XHTML tienen que tener un único elemento raíz de todo el documento, el elemento `<html>`. Los documentos HTML pueden no tenerlo.

HTML ✓	<pre><!DOCTYPE html> <html> <head> <meta charset="utf-8"> <title>HTML</title> </head> <body> <!-- contenido de la página --> </body> </html></pre>
XHTML ✗	
HTML ✓	<pre><!DOCTYPE html> <head> <meta charset="utf-8"> <title>HTML</title> </head> <body> <!-- contenido de la página --> </body></pre>
XHTML ✗	
HTML ✗	<pre><?xml version="1.0"?> <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd"> <html xmlns="http://www.w3.org/1999/xhtml"> <head> <meta http-equiv="Content-Type" content="text/html; charset=utf-8" /> <title>XHTML</title> </head> <body> <!-- contenido de la página --> </body> </html></pre>
XHTML ✓	