

1. Introducción

1.1 ¿Qué es HTML?

Definiéndolo de forma sencilla, **"HTML es el lenguaje que se utiliza para crear las páginas web a las que se accede mediante internet"**. Más concretamente, HTML es el lenguaje con el que se "escriben" la mayoría de páginas web.

Los diseñadores utilizan el lenguaje HTML para crear páginas web, los programas que se emplean generan páginas escritas en HTML y los navegadores que utilizamos (por ejemplo Google Chrome, Opera o Mozilla Firefox) muestran las páginas web después de leer e interpretar su contenido HTML.

Aunque HTML es un lenguaje que utilizan los ordenadores y los programas de diseño de páginas web, es muy fácil de entender y escribir por parte de las personas. En realidad, HTML son las siglas de *HyperText Markup Language* y más adelante se verá el significado de cada una de estas palabras. El lenguaje HTML es un estándar reconocido en todo el mundo y cuyas normas define un organismo sin ánimo de lucro llamado [World Wide Web Consortium](#), más conocido como **W3C**. Como se trata de un estándar reconocido por todas las empresas relacionadas con el mundo de internet, una misma página escrita en HTML se visualizará de forma muy similar en cualquier navegador bajo distintos sistemas operativos.

El propio **W3C** define el lenguaje HTML como *"un lenguaje reconocido universalmente y que permite publicar información de forma global"*. Desde su creación, el lenguaje HTML ha pasado de ser un lenguaje utilizado exclusivamente para crear documentos electrónicos a emplearse en diversas aplicaciones electrónicas especializadas como buscadores, tiendas "en línea" y banca electrónica.

1.2. Breve historia de HTML

La historia completa de HTML es algo larga, por lo que se muestra resumida a partir de la información que se puede encontrar en [Wikipedia](#).

El origen de HTML se remonta a 1980, cuando el físico [Tim Berners-Lee](#), investigador del [CERN](#) (Organización Europea para la Investigación Nuclear) propuso un nuevo sistema de "hipertexto" para compartir documentos.

Los sistemas de hipertexto habían sido desarrollados años antes. En el ámbito de la informática, el hipertexto permite que los usuarios accedan a la información relacionada con los documentos electrónicos que visualizan. En cierta manera, los primitivos sistemas de hipertexto podrían asimilarse a los enlaces de las páginas web actuales.

Tras finalizar el desarrollo de su sistema, Tim Berners-Lee lo presentó a una convocatoria organizada para desarrollar un sistema de hipertexto para internet. Después de unir sus fuerzas con el ingeniero de sistemas Robert Cailliau, presentaron la propuesta ganadora llamada *WorldWideWeb* (W3). El primer documento formal con la descripción de HTML se publicó en 1991 bajo el nombre *"HTML Tags"* (Etiquetas HTML) y todavía hoy puede ser [consultado en línea](#) a modo de reliquia informática.

La primera propuesta oficial para convertir HTML en un estándar se realizó en 1993 por parte del

organismo [IETF](#) (*Internet Engineering Task Force*). Aunque se consiguieron avances significativos (en esta época se definieron las etiquetas para imágenes, tablas y formularios) ninguna de las dos propuestas de estándar, llamadas HTML y HTML+ consiguieron convertirse en estándar oficial.

En 1995, el organismo IETF organizó un grupo de trabajo de HTML y el 22 de septiembre publicaron el estándar HTML 2.0. A pesar de su nombre, HTML 2.0 es el primer estándar oficial de HTML.

A partir de 1996, los estándares de HTML los publica otro organismo de estandarización, el W3C. La versión HTML 3.2 se publicó el 14 de Enero de 1997 y es la primera recomendación de HTML publicada por el W3C. Esta revisión incorpora los últimos avances de las páginas web desarrolladas hasta 1996, como *applets* de Java y texto que fluye alrededor de las imágenes.

HTML 4.0 se publicó el 24 de Abril de 1998 (versión corregida de la publicación del 18 de Diciembre de 1997) y supuso un gran salto desde las versiones anteriores. Entre sus novedades más destacadas se encuentran las hojas de estilos CSS, la posibilidad de incluir pequeños programas o *scripts* en las páginas web, mejora de la accesibilidad de las páginas diseñadas, tablas complejas y mejoras en los formularios.

La última especificación oficial de HTML se publicó el 24 de diciembre de 1999 y se denomina HTML 4.01. Se trata de una revisión y actualización de la versión HTML 4.0, por lo que no incluye novedades significativas.

Desde la publicación de HTML 4.01, la actividad de estandarización de HTML se detuvo y el W3C se centró en el desarrollo del estándar XHTML. Por este motivo, en el año 2004, las empresas Apple, Mozilla y Opera mostraron su preocupación por la falta de interés del W3C en HTML y decidieron organizarse en una nueva asociación llamada [WHATWG](#) (*Web Hypertext Application Technology Working Group*). La actividad actual del WHATWG se centra en el estándar HTML5, cuyo primer borrador oficial se publicó el 22 de enero de 2008. Debido a la fuerza de las empresas que forman el grupo WHATWG y a la publicación de los borradores de HTML 5.0, en marzo de 2007 el W3C decidió retomar la actividad estandarizadora de HTML.

HTML5 ya es un estándar (recomendado), el consorcio W3C, con el inventor de la Web Sir Tim Berners-Lee, presentó el 27 de octubre de 2014 la versión final. Su intención es que se construya lo que se ha llamado Plataforma Web abierta, donde el HTML5, junto a Javascript y CSS3, se podrá utilizar para el desarrollo de aplicaciones multiplataforma (Linux, Windows, Android, iOS..).

Como parte de la estrategia para vincular a más gente en el proceso actual, el director ejecutivo del Consorcio, el Dr. [Jeff Jaffe](#) (MIT) ha publicado un texto acerca de las prioridades del mismo bajo el título de ‘[Fundamentos de Aplicación](#)’ que sostienen todo el trabajo orientado hacia la plataforma actual y de “nueva generación”:

- Seguridad y privacidad, todo lo relacionado con autenticaciones, encriptación, protección de identidad y actividad en línea.
- Diseño y desarrollo de la web, en cuanto a estilo, formato, gráficos, animación y tipografía.
- Interacción con distintos equipos como sistemas de sensores y Bluetooth.
- Ciclo de uso de aplicación para administración de tareas fuera de conexión y sincronización.
- Medios y comunicaciones en tiempo real, para efectos, por ejemplo, de transmisiones en vivo (*streaming*).
- Desempeño y afinación de la capacidad y precisión en la respuesta y descarga de sitios web con sus funciones.
- Usabilidad y accesibilidad, para un web internacional, multilingüe y de acceso para personas con distintas discapacidades.
- Servicios como pagos y web social.

De forma paralela a su actividad con HTML, W3C ha continuado con la estandarización de XHTML, una versión avanzada de HTML y basada en XML. La primera versión de XHTML se denomina XHTML 1.0 y se publicó el 26 de Enero de 2000 (y posteriormente se revisó el uno de Agosto de 2002).

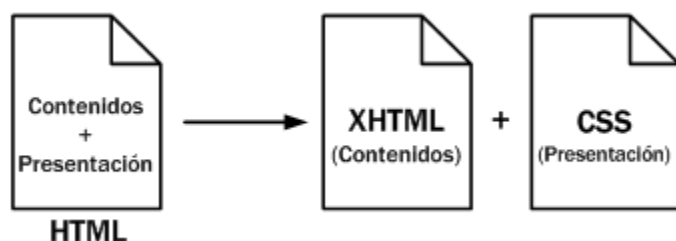
XHTML 1.0 es una adaptación de HTML 4.01 al lenguaje XML por lo que mantiene casi todas sus etiquetas y características, pero añade algunas restricciones y elementos propios de XML. La versión XHTML 1.1 ya ha sido publicada en forma de borrador y pretende modularizar.

1.3. HTML y CSS

Originalmente, las páginas HTML sólo incluían información sobre sus contenidos de texto e imágenes. Con el desarrollo del estándar HTML, empezaron a incluir también información sobre el aspecto de sus contenidos: tipos de letra, colores y márgenes.

La posterior aparición de tecnologías como *JavaScript*, provocaron que las páginas HTML también incluyeran el código de las aplicaciones (*scripts*) que se utilizan para crear páginas web dinámicas.

Incluir en una misma página HTML los contenidos, el diseño y la programación complica en exceso su mantenimiento. Normalmente, los contenidos y el diseño de la página web son responsabilidad de diferentes personas, por lo que es adecuado separarlos. CSS es el mecanismo que permite separar los contenidos definidos mediante XHTML y el aspecto que deben presentar esos contenidos:



Otra ventaja de la separación de los contenidos y su presentación es que los documentos XHTML creados son más flexibles, ya que se adaptan mejor a las diferentes plataformas: pantallas de ordenador, pantallas de dispositivos móviles, impresoras y dispositivos utilizados por personas discapacitadas.

De esta forma, utilizando exclusivamente XHTML se crean páginas web "feas" pero correctas. Aplicando CSS, se pueden crear páginas "bonitas" a partir de las páginas XHTML correctas.

2. Características básicas

2.1 Lenguajes de etiquetas

Uno de los retos iniciales a los que se tuvo que enfrentar la informática fue el de cómo almacenar la información en los archivos digitales. La solución que se emplea para guardar la información con formato es sencilla: el archivo electrónico almacena tanto los contenidos como la información sobre el formato de esos contenidos. Si por ejemplo se quiere dividir el texto en párrafos y se desea dar especial importancia a algunas palabras, se podría indicar de la siguiente manera:

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

El principio de un párrafo se indica mediante la palabra `<parrafo>` y el final de un párrafo se indica mediante la palabra `</parrafo>`. De la misma manera, para enfatizar ciertas palabras del texto, se encierran entre `<importante>` e `</importante>`. El proceso de indicar las diferentes partes que componen la información se denomina marcar (*markup* en inglés). Cada una de las palabras que se emplean para marcar el inicio y el final de una sección se denominan etiquetas.

Aunque existen algunas excepciones, en general 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 `>`.

Así, la estructura típica de las etiquetas HTML es:

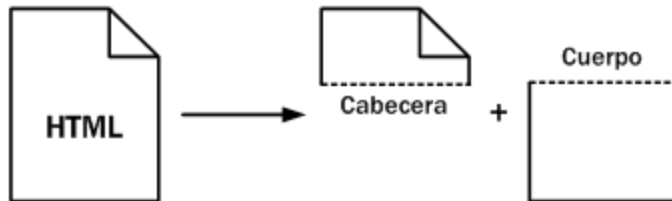
```
<nombre_etiqueta> ... </nombre_etiqueta>
```

HTML es un lenguaje de etiquetas (también llamado lenguaje de marcas) y las páginas web habituales están formadas por cientos o miles de pares de etiquetas. De hecho, las letras "ML" de la sigla HTML significan *markup language*, que es como se denominan en inglés a los lenguajes de marcas. Además de HTML, existen muchos otros lenguajes de etiquetas como XML, SGML, DocBook y MathML.

La principal ventaja de los lenguajes de etiquetas es que son muy sencillos de leer y escribir por parte de las personas y de los sistemas electrónicos. La principal desventaja es que pueden aumentar mucho el tamaño del documento, por lo que en general se utilizan etiquetas con nombres muy cortos.

2.2. El primer documento HTML

Las páginas HTML se dividen en dos partes: la cabecera y el cuerpo. La cabecera incluye información sobre la propia página, como por ejemplo su título y su idioma, que no se visualiza. El cuerpo de la página incluye todos sus contenidos, como párrafos de texto e imágenes.



El cuerpo (llamado *body* en inglés) contiene todo lo que el usuario ve en su pantalla y la cabecera (llamada *head* en inglés) contiene todo lo que no se ve (con la única excepción del título de la página, que los navegadores muestran como título de sus ventanas).

A continuación se muestra el código HTML de una página web muy sencilla:

```
<html>
<head>
<title>Mi primer documento HTML</title>
</head>

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

</html>
```

Para probar este primer ejemplo, se ha de hacer lo siguiente:

Abre un editor de archivos de texto y crea un archivo nuevo

Copia el código HTML mostrado anteriormente y ponlo tal cual en el archivo que has creado

Guarda el archivo con el nombre que quieras, pero con la extensión html

Llama el fichero desde un navegador de internet (no es necesario llevarlo a un servidor).

3. Cómo lograr un sitio web atractivo

Si tus estadísticas están cayendo y cada vez son más las personas que no regresan a tu sitio, es obvio que existe algo que no funciona. Desarrollar un sitio atractivo que capte una y otra vez a los lectores no es una tarea sencilla. Las causas de la mala aceptación pueden ser muchas, lo importante es analizar el sitio objetivamente y valorar qué factores están incidiendo en la percepción de los visitantes:

1. **La saturación visual de textos e imágenes.** Es esencial que el diseño del sitio web posea una estructura ordenada de forma que no presente una imagen caótica. Se ha comprobado que los sitios particularmente desorganizados provocan estrés en los visitantes por lo que, probablemente, estos no querrán regresar. Cuando se trata de diseño, recuerda que **menos es más**.

2. **Una navegación confusa.** Los usuarios de internet suelen premiar la simplicidad. Y el mejor ejemplo de ello es Apple. Si tu sitio web cuenta con menús extensos y con múltiples enlaces, esto provoca la sensación de desorden y al visitante le será difícil moverse dentro del sitio web y comprender su organización. Recuerda además, que la organización da sensación de control y seguridad por lo que tus visitantes querrán volver.

3. **Incompatibilidad con navegadores y dispositivos móviles.** Aunque no seas un programador experto, debes asegurarte de que tu sitio web tenga una visibilidad aceptable en los principales navegadores que se usan en la actualidad así como que posea una versión para los teléfonos inteligentes, tabletas y lectores de libros. De esta forma no te estarás cerrando a ninguna posibilidad. Debes recordar que a veces los navegadores producen superposiciones de textos e imágenes o provocan que el diseño se corra obligando a realizar desplazamientos muy largos. Por ende, lo mínimo que puedes hacer es revisar la funcionalidad de tu sitio web en los diferentes navegadores y dispositivos.

4. **Publicidad excesiva.** Cuando los sitios web están tan llenos de publicidad que cuesta encontrar el contenido y cuando se usan ventanas *pop up* que molestan al visitante, es muy probable que este no vuelva. Por supuesto, la solución no está en eliminar la publicidad sino en insertarla armónicamente con el diseño web.

5. **Lentitud al cargar los contenidos.** Con la inmediatez de internet, no hay nada más exasperante que un sitio web que se demore en cargar. Esto puede suceder por varios motivos: un mal proveedor de alojamiento o un sitio web con imágenes demasiado pesadas. De una forma u otra, ten en cuenta que 10 segundos ya se considera un sitio lento en cargar. En [Pingdom](#) podrás verificar cuánto tarda un sitio web.

6. **Usar plantillas populares.** Existen abundantes plantillas populares, muchas de ellas las ofrecen las mismas plataformas como Blogger y Wordpress. El uso de ellas puede ser contraproducente ya que anulan toda la identidad del sitio web. Recuerda que en un mundo tan amplio como internet, diferenciarse es esencial.

7. **Contenido poco original.** En la red ya existen cientos de sitios que copian los contenidos de otros. En la misma medida en que tus visitantes comiencen a percatarse de que copias siempre de los mismos sitios web, optarán por visitar directamente a estos.

8. **Ausencia de una temática principal.** Independientemente del número de tópicos que se pueden abordar, es importante que el visitante aprecie una consistencia temática. De esta forma se puede crear una identidad que te diferencie del resto de los sitios de la web.

9. **Web no accesible.** La accesibilidad web significa que personas con algún tipo de discapacidad van a poder hacer uso de la web. Se está haciendo referencia a un diseño que va a permitir que estas personas puedan percibir, entender, navegar e interactuar con la web, aportando a su vez contenidos. La accesibilidad también beneficia a otras personas, incluyendo personas de edad avanzada que han visto mermadas sus habilidades. Se rige por las normas: Norma UNE 139803:2012, WCAG 2.0, Norma EN 301 549 y WAI-ARIA.