UNIDAD 1. GENERALIDADES HTML5

# GENERALIDADES HTML5

Este bloque trata de dar una idea general de lo que es HTML5, en qué se basa, cuáles son sus predecesores y cuáles son sus ventajas sobre otras versiones de HTML.

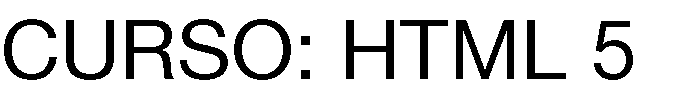
Para entender el funcionamiento básico de HTML5 es importante conocer los fundamentos de los lenguajes de marcas y su evolución desde sus primeras versiones allá por los años 60 ya que nos va a permitir introducirnos en el mundo de la creación de páginas web.

HTML5 simplifica y mejora la creación de documentos HTML, eliminando etiquetas obsoletas, apoyándose en las hojas de estilo para mejorar su aspecto, en javascript para ganar dinamismo y funcionalidad y en sus nuevos elementos para no tener que usar programas externos o para crear una estructura más coherente dentro de una página web.

Ante las innovaciones que aporta HTML5 es interesante saber el grado de compatibilidad que tienen los navegadores a la hora de tratar con los nuevos elementos HTML5 ya que esto nos permitirá utilizar esos elementos para que no haya problema a la hora de verlos en cualquier navegador.

También hay que ver que HTML5 va más allá de lo que es un lenguaje de marcas y verlo como una agrupación de especificaciones que están marcando el desarrollo web.

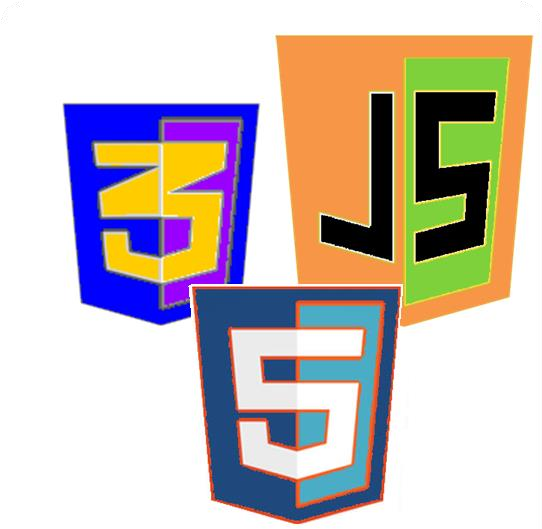
En definitiva, tal y como lo describe el consorcio W3C, **HTML5 es el futuro del contenido.**



# DEFINICIÓN DE HTML5

HTML5 es lo que se denomina un **lenguaje de marcas** que, combinado con **CSS3**, **JavaScript** y algunas tecnologías de diseño, permite la creación de páginas web.

Es la quinta revisión importante del lenguaje básico de Internet denominado HTML.

Cuando creamos una página web debemos saber que esta no es ni más ni menos que un documento de texto compuesto por contenidos o textos que manejan los elementos HTML, los cuales se rigen por las normas del lenguaje HTML, y que se visualiza a través de una aplicación denominada navegador (Google Chrome, Internet Explorer, Ópera, Mozilla Firefox, etc.) que interpreta su contenido HTML.

El lenguaje HTML es un estándar cuyas normas, a partir de su versión 3.0, son definidas por un organismo sin ánimo de

lucro denominado World Wide Web Consortium, más conocido como **W3C**.

El propio W3C define el lenguaje HTML como "un lenguaje reconocido universalmente y que permite publicar información de forma global".

Esta definición debería implicar que una misma página web se vería igual a través de un navegador u otro, pero esto no ocurre ya que los diferentes navegadores interpretan de forma muy desigual el lenguaje HTML de una misma página web dejando la visualización del contenido más en manos del navegador que de su propio diseño.

El estándar HTML5 tratará en el futuro de paliar este problema y todos los navegadores lo deberán interpretar de la misma manera.



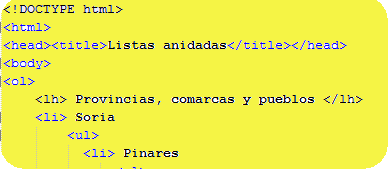
# LENGUAJES DE MARCAS

#### ¿Qué son?

Cuando escribimos un documento y queremos resaltar un carácter, palabra o frase lo que hacemos normalmente es subrayarlo, escribirlo más fuerte o usar caracteres más grandes para que destaque del resto del texto, es decir, estamos aplicando un **formato** concreto al texto con el objeto de distinguirlo de la información que está a su alrededor.

Utilizando el párrafo anterior a modo de ejemplo, la palabra **formato**, destaca del resto por queremos llamar la atención sobre él, por eso la hemos puesto en negrita y con un tipo de letra mayor que la del resto del documento, es decir, le hemos dado un determinado formato.

Un lenguaje de marcas define un conjunto de marcas o etiquetas y la estructura que debe tener un documento para poder aplicar dichas marcas.



Una "marca" o “etiqueta” es una señal colocada dentro de un documento que delimita una parte de él a la que aplica un determinado formato en base a la función que tiene asignada dicha marca en el lenguaje.

*Un lenguaje de marcas no es un lenguaje de programación*



*Un lenguaje de marcas se puede combinar con un lenguaje de programación para aportar nuevas funcionalidades y dar más calidad a la página web*

#### Ejemplos:

|  |  |
| --- | --- |
| **Lenguaje de marcas** | **Texto que se visualiza** |
| Hola Mundo. Adiós mundo. |  |
| <b> Hola Mundo.</b> |  |
| <font size=”7”> Adiós mundo  </font> |  |

En el primer ejemplo se ve el texto tal y como se escribe ya que dicho texto no está dentro de una etiqueta. En el segundo ejemplo, la etiqueta <b> pone en negrita el texto que hay hasta la etiqueta de cierre </b> y, en el tercer ejemplo, <font size=”7”> pone el texto hasta la etiqueta de cierre </font> a un mayor tamaño

#### Tipos

Se suelen diferenciar tres tipos de lenguajes de marcas según el tipo de marcas que definen, aunque en un mismo documento puede haber marcas de los tres tipos. Estos lenguajes son:

* + - * *Lenguaje de marcas de presentación*: Es aquel cuyas marcas definen el formato del texto que engloban lo que permite modificar el aspecto estético de un documento, aunque no sea suficiente para el procesamiento de información.
      * *Lenguaje de marcas de procedimiento*: Es aquel cuyas marcas son una serie de órdenes normalmente de tipo tipográfico para la presentación de textos.
      * *Lenguaje de marcas descriptivas o semánticas*: Es aquel cuyas marcas identifican bloques dentro del documento y describen su contenido sin entrar en aspectos de formato.

*HTML es un lenguaje de marcas descriptivas o semánticas*



#### Características

* + - * *Texto Plano*

Los archivos de texto plano son aquellos que están compuestos únicamente por caracteres de texto, a diferencia de los archivos binarios que pueden contener imágenes, sonido, archivos comprimidos, programas compilados, etc.

Una de las principales ventajas de los archivos de texto plano es que pueden ser interpretados directamente por un simple editor de texto, a diferencia de los binarios que necesitan software específico (visores, descompresores, compiladores, etc.) Esta característica hace que los documentos sean independientes del sistema operativo o programa con el que fueron creados, esto facilita la interoperabilidad, que constituye una importante ventaja para el intercambio de información en Internet.

* + - * *Integración*

Las instrucciones de marcado se mezclan con el propio contenido. El texto entre las marcas es el propio contenido del documento.

* + - * *Independencia del dispositivo*

Inicialmente los lenguajes de marcas se idearon para visualizar documentos de texto, pero progresivamente se han empezado a utilizar en muchas otras áreas como gráficos vectoriales, sindicación de contenidos, notación científica, interfaces de usuario, síntesis de voz, etc.

* + - * *Flexibilidad*

Los lenguajes de marcas se pueden combinar en el mismo archivo con otros lenguajes, como HTML con PHP y JavaScript. Incluso hay etiquetas específicas para ello como es <script>.

#### Ejemplos de lenguajes de marcas

* + - * *SGML* (Standard Generalized Markup Language): Al igual que su antecesor GML, SGML o ISO 8879 es un lenguaje de marcas genérico, contiene las normas para generar otros lenguajes de marcas sin tener en cuenta su codificación interna por lo que se puede denominar que es un metalenguaje del que se derivan otros lenguajes más especializados y menos generalistas. Esto se hace posible a través del uso de una definición de tipo de documento (DTD, Document Type Definition), en la que se

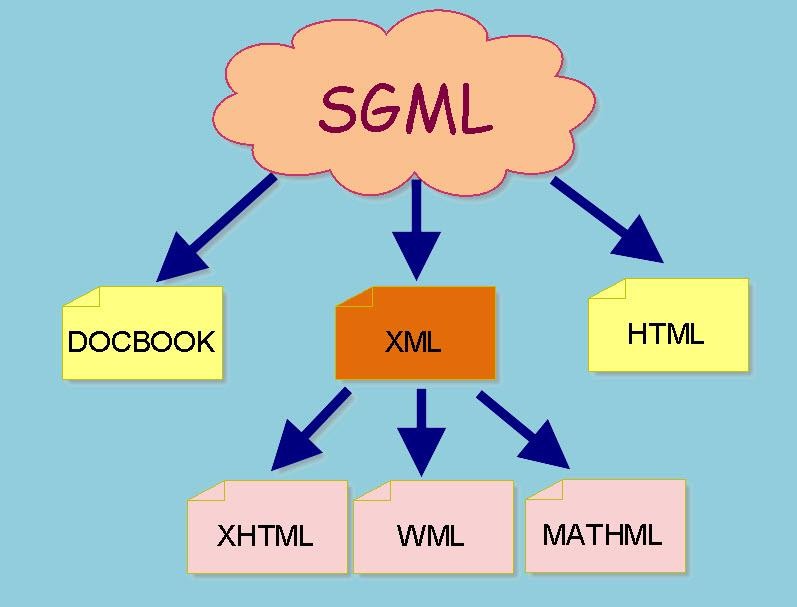
especifican los elementos, atributos y su orden que se emplean en los documentos que cumplen un determinado lenguaje.



* + - * *HTML*: Basado en SGML, HTML o Hypertextual Markup Language (lenguaje de marcas de hipertexto), es un estándar para la creación de páginas web que define una estructura básica (todos los documentos HTML tienen una cabecera <HEAD> y un cuerpo <BODY>) y un código genérico, denominado HTML, que indica los elementos y su funcionalidad de los mismos para todas las versiones, aunque cada una de ellas tiene su propio DTD con los componentes y normas que las diferencian.

*El organismo W3C se encarga de mantener las DTD de los lenguajes de marcas más usados como las versiones de HTML a excepción de HTML5 que no tiene.*

* + - * *XML*: Subconjunto de SGML adaptado a las necesidades de intercambio de datos entre diversas aplicaciones, XML o Extensible Markup Language (lenguajes de marcas extensible) no es propiamente un lenguaje de marcas (no se usa para hacer páginas web) sino que es un metalenguaje en el que se basan otros lenguajes de marcas ya que se pueden definir etiquetas para determinadas funciones.
      * *XHTML*: Es nueva adaptación de HTML con arreglo a las normas XML lo que obliga a que los documentos XHTML, Extensible HyperText Markup Language, deban estar “bien formados”, es decir, deben cumplir estrictamente la especificación XML, cosa que no sucede con los documentos HTML, lo que hace que los navegadores los visualicen más fidedignamente.



# HISTORIA DE HTML5

IBM desarrolló en 1986 un lenguaje de marcas denominado GML (Generalized Markup Language) que pretendía resolver los problemas de compatibilidad que había entra aplicaciones debido a que cada aplicación utilizaba sus propias etiquetas o marcas para describir sus diferentes elementos y era prácticamente imposible el intercambio de información.

GML usa marcas genéricas que permiten distinguir el contenido del documento del aspecto y estructura del mismo.

Ese mismo año GML pasó a ser de la organización ISO convirtiéndose en el lenguaje SGML (ISO 8879), Standard Generalized Markup Language, con las características de ser un lenguaje de código abierto y de libre disposición. A partir de este momento se puede hablar de SGML como metalenguaje, es decir, lenguaje que sirve de base a otros lenguajes, en este caso, de marcas debido a que tiene un grupo de elementos y unas reglas que marcan la utilización de los elementos anteriores y sus atributos.

|  |  |
| --- | --- |
| **AÑO** | **Lenguaje** |
| -1986 | GML |
| 1986 | SGML |
| 1991 | HTML |
| 1996 | CSS, JavaScript |
| 1998 | HTML4 |
| 2000 | XHTML 1.0 |
| 2003 | XHTML 2.0 |
| 2009 | HTML5 |

En 1991 nace HTML (HiperText Markup Language) a partir de SGML con la finalidad de ser el lenguaje que se emplee para la elaboración de páginas web. En sus versiones finales como HTML 4.01 se agregan hojas de estilo CSS para los aspectos estéticos y javascript para darle más dinamismo.

Después de HTML 4.01, el consorcio W3C, comunidad internacional que desarrolla estándares como HTML, desarrolló el lenguaje XHMTL 1.0 que es mucho más rígido que HTML (es obligatorio escribir en minúsculas las etiquetas) pero que intentaba que todos los navegadores vieran igual el mismo documento HTML, cosa que no pasa como HTML 4.01.

A continuación, se desarrolló el lenguaje XHTML 2.0 que modifica y mejora el lenguaje XHTML 1.0

Seguidamente un grupo de fabricantes de navegadores y desarrolladores web crearon una nueva especificación de HTML superior a HTML 4.01 orientada a crear un nuevo tipo de aplicaciones web denominada HTML5 y se abandonó el desarrollo de XHTML 2.0.

Es conveniente usar las versiones más modernas de navegadores que reconozcan los nuevos elementos HTML5 para poder sacar rendimiento a este lenguaje.

# Compatibilidad de navegadores con HTML5

Los pasos que seguimos cuando queremos ver una página web creada por nosotros son:

* Creación del documento HTML con un editor normal o con un editor web siguiendo las normas y usando los elementos de la versión HTML que queramos utilizar en dicho documento.
* Visualización del documento con un navegador web como puede ser Google Chrome, Mozilla Firefox o Internet Explorer, entre otros.

Pero cuál es nuestra sorpresa al ver que la misma página web no se ve igual en distintos navegadores. Es más, si nos referimos al nuevo estándar HTML5, vemos que hay

elementos que ni tan siquiera los reconocen algunos navegadores.

Ante esta perspectiva lo único que podemos hacer es consultar en

páginas de referencia en Internet el grado de compatibilidad con HTML5 que tiene un navegador, es decir, que nos indique los elementos y características HTML5 que soporta dicho navegador.

La razón de hacer esto se debe a que debemos conocer cuáles son los elementos que se reconocen en los navegadores más utilizados por los usuarios y usarlos en nuestras páginas. Esto nos dará la certeza de que nuestra página web se visualizará sin problemas en cualquier navegador.

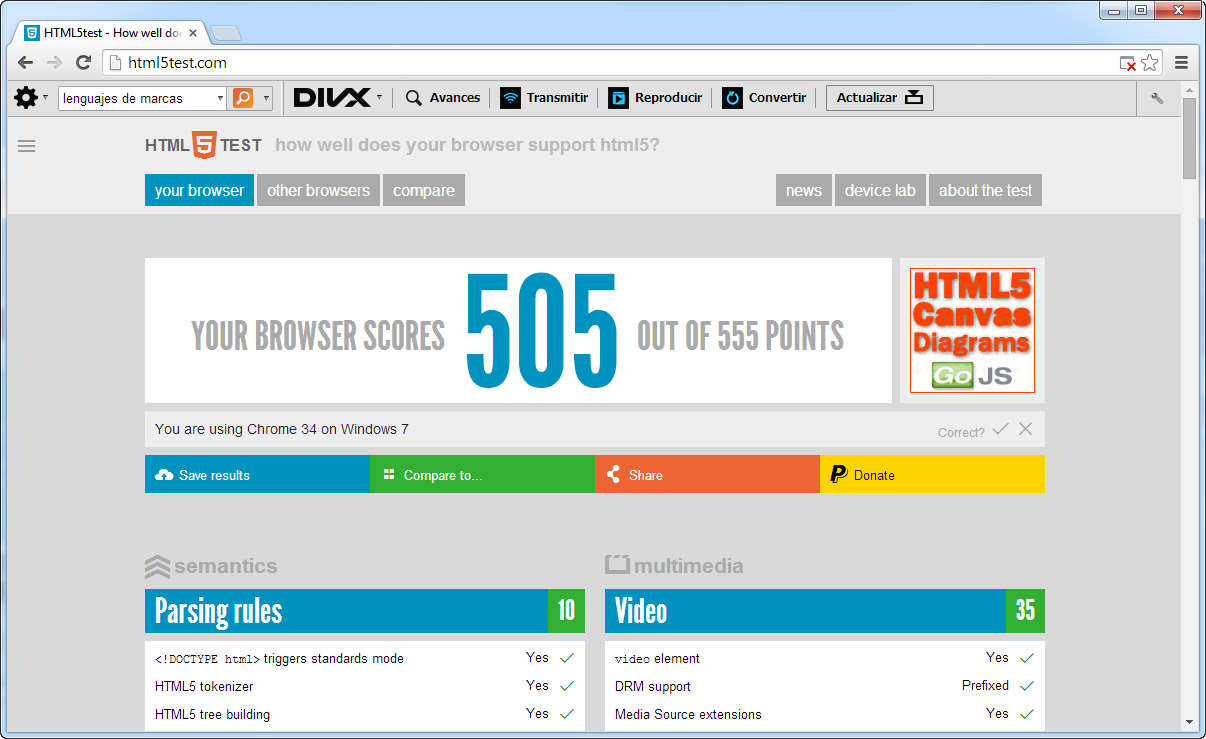
Otro problema añadido es la controversia sobre cuáles son los navegadores más utilizados en la red ya que hay estadísticas muy dispares sobre el tema, pero podemos optar por informarnos sobre las últimas versiones de Google Chrome, Internet Explorer, Mozilla Firefox y Ópera.

Una de estas páginas de consulta es [http://html5test.com](http://html5test.com/) que permite conocer la compatibilidad del navegador con el que se visualiza dicha página y tiene comparativas entre los principales navegadores que se usan incluidos los de Tablet y móviles.



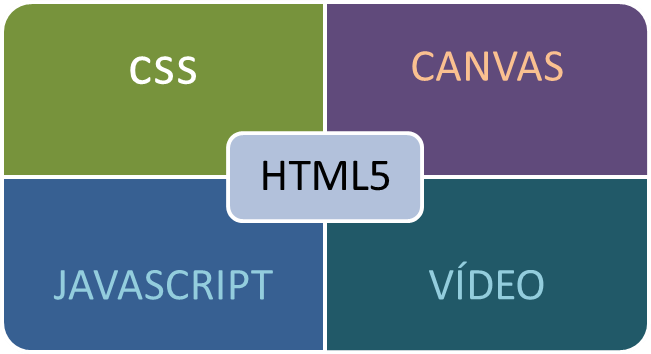
*El navegador Google Chrome en cualquiera de sus versiones es el que mejor reconoce los nuevos elementos y características HTML5 además de ser, posiblemente, el más utilizado por los usuarios.*

**Ejercicio**: Comprueba el grado de compatibilidad con html5 que tienen tus navegadores y compáralos a través de la página [http://html5test.com](http://html5test.com/)



# VENTAJAS DE HTML5

* Mayor compatibilidad entre los navegadores ya que el aspecto de una página web en uno u otro navegador no difiere tanto como cuando se visualizaban páginas HTML 4 o XHTML.
* Mayor rapidez a la hora de cargar la página debido a la simplificación de las etiquetas.
* Nuevas etiquetas que permiten crear una estructura bastante intuitiva de un documento HTML.
* Inserción de audio y vídeo de forma directa sin necesidad de emplear ningún plug- in.
* Permite la geolocalización de un usuario.
* Incorporación de más efectos visuales a través de la etiqueta canvas.
* El aspecto estético del documento se basa casi por completo en el uso de hojas de estilo CSS eliminando o reduciendo a la mínima expresión el uso de etiquetas de formato.
* La incorporación de nuevas capacidades Javascript.



# NOVEDADES

#### Nuevas etiquetas estructurales

La estructura normal de una página web tiene un armazón bastante concreto basado en una cabecera, menús, cuerpos, pies, etc. Con HTML5 se usan etiquetas específicas (denominadas semánticas porque aportan información relevante) para definir todos estos elementos. Su uso no es obligatorio, pero sí bastante descriptivo ya que un vistazo sobre el código html nos indica lo que hay en cada sección. Las etiquetas html4 <div> y <span> empleadas para definir partes de una página prácticamente no se usan ya que <header>, <nav>, <section> y otras etiquetas nuevas de html5 son las encargadas de definir las partes de una página web. Estos son los elementos:

* **Section:** Representa un bloque general dentro de un documento, es decir, un bloque de una misma temática. Puede contener subsecciones o artículos.
* **article** representa un contenido específico en un documento.
* **aside** representa un contenido complementario o muy poco relacionado con el resto de la página.
* **header** representa la cabecera de una página html o de una sección.
* **footer** representa el pie de una página o sección, con información acerca de la misma que va al final de la página o sección.
* **nav** zona de navegación con enlaces dispuestos de diferentes maneras.
* **hgroup: Sirve para el agrupamiento de titulares.**

#### Nuevas etiquetas para video, audio y tratamiento de gráficos

Las etiquetas <video> y <audio> permiten la inserción de audio y vídeo en una página html5 de manera nativa sin depender de los plug-ins o código que necesitan los navegadores para saber qué hacer con un determinado archivo.

La etiqueta <canvas> permite trabajar con gráficos y animaciones.

La etiqueta <embed> sirve para manejar contenido incrustado, pero no nativo, sino ejecutado por plugins como el de Flash.

#### Nuevas APIS

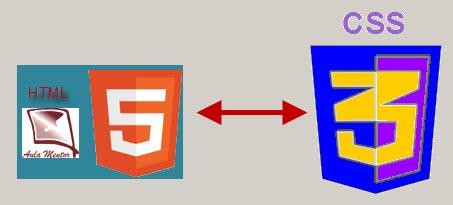
Una interfaz de programación de aplicaciones (API) es un conjunto de instrucciones, funciones y normas de programación para acceder a una aplicación de software. Uno de los objetivos principales de una API es el de normalizar la comunicación entre aplicaciones.

Las APIs son un aspecto muy importante dentro del entorno de HTML5 y hay una serie de ellas que conviene conocer, como Web Storage, Microdata o Geolocation (que permite conocer el punto geográfico desde el cual se conecta el navegador a Internet), entre otras y, en el futuro, se van a incorporar varias APIS nuevas que

van a permitir trabajar con Bases de datos en local, drag & drop (arrastrar y soltar) lo cual va a hacer de html5 un lenguaje mucho más potente.

* + **Relación muy fuerte con hojas de estilos CSS3**

Los aspectos estéticos y de formato de una página html5 recaen casi completamente sobre el uso de Hojas de Estilo en Cascada (CSS) lo que facilita separar contenidos de formatos de presentación. De hecho, el estándar 3.0 de CSS ha ido evolucionando en paralelo con html5, tanto que hay algunos componentes CSS3 que se confunden con elementos html5.

Por esta razón la inmensa mayoría de las etiquetas html 4 de formato pasan a quedar obsoletas y a no ser utilizadas con lo que se simplifica el número de etiquetas HTmL5.

* + **Mejores formularios**

Un formulario es una sección compuesta de controles a través de los cuales se recopila información introducida por el usuario para luego ser utilizada.

HTML5 añade características nuevas para la realización de formularios mejorando sustancialmente los de la anterior versión de HTML o HTML4.01.

La principal novedad es la validación de datos por parte del navegador o del lado cliente lo que no hará necesario el empleo de programas auxiliares y se simplifica notablemente tanto la creación como el uso de formularios. Para ello incluye nuevos campos de entrada de datos y atributos para otros controles.



Un ejemplo que constata todo lo anterior es el elemento input el cual ha sido simplificado en su sintaxis, pero ampliado en funcionalidades.

No todos los navegadores permiten la validación del lado cliente. En el futuro si lo deberían de poder hacer.

# CONCLUSIÓN

HTML5 representa el futuro de la creación de páginas web, pero debemos saber que, aunque simplifique y estructure mejor el código HTML, agregue nuevos atributos y etiquetas y elimine aquellas que están obsoletas, no es solo una nueva versión de un lenguaje de marcas HTML sino una agrupación de diversas especificaciones que están influyendo en el desarrollo web.

A su vez, HTML5 pretende proporcionar una plataforma con la que desarrollar aplicaciones web muy parecidas a las aplicaciones de escritorio, es decir, cercano a lo que se denomina Web 2.0. (uso de aplicaciones de web en vez de locales y que la propia web sea la plataforma donde se integran estas aplicaciones). Para ello se están creando APIs que permitan trabajar con cualquiera de los elementos de la página sin necesidad de usar programas adicionales lo que debería generar un estándar que todos los navegadores lo reconocieran de la misma manera y desaparecieran los problemas de compatibilidad entre las funcionalidades de este estándar y los navegadroes.

Por eso el consorcio W3C habla de HTML5 como El futuro del contenido web.

