

H.3. Html5: el nuevo estándar básico de la Web

Por **Jorge Franganillo**

6 septiembre 2010

Franganillo, Jorge. "Html5: el nuevo estándar básico de la Web".
Anuario ThinkEPI, 2011, v. 5, pp. 261-265.



Resumen: A finales de 2009 el W3C detuvo el diseño de xhtml 2 y se concentró en definir html5. Desde entonces ha habido cierta confusión y acalorados debates sobre cómo debería ser el lenguaje destinado a sustentar el futuro desarrollo de la Web. Html5 se ha concebido con el propósito de simplificar el trabajo de los diseñadores de webs y mejorar el rendimiento de las páginas, especialmente en dispositivos móviles. Se analiza el último borrador de la especificación html5 y se comentan los aspectos más significativos: la sintaxis, el DOM, la presentación y la recogida de información, y el procesamiento de contenido multimedia.

Palabras clave: Html5, Estándares web, Lenguajes de etiquetado, DOM, Flash, Navegadores web.

Title: *Html5: the new basic standard of the Web*

Abstract: In late 2009 the W3C stopped developing xhtml2 and focused on defining html5 instead. Since then, there has been some confusion and heated debates about the proper markup language for future web development. Html5 was conceived with the aim of simplifying the designers' work and improving the sites' performance, especially on mobile devices. This paper analyzes the final draft of the html5 specification and discusses the most significant issues: syntax, DOM, information gathering and presentation, and multimedia processing.

Keywords: Html5, Web standards, Mark-up languages, DOM, Flash, Web browsers.

EL LENGUAJE html es uno de los puntales de la Web. Desde hace más de dos décadas ejerce una función primordial en el proceso de distribuir información a través de internet.

En todo ese tiempo ha experimentado algunos cambios en respuesta a las necesidades de los productores de contenido y de los usuarios, y ahora se encuentra a punto para el cambio siguiente: el *World Wide Web Consortium (W3C)* está esbozando html5.

Esta nueva versión del lenguaje básico de la Web proporciona mecanismos para simplificar el trabajo y facilitar la inclusión de elementos multimedia. El principal criterio de diseño de html5 ha sido el de resolver problemas prácticos, y con este objetivo adopta soluciones dirigidas a facilitar el trabajo en situaciones reales.

"A mediados de 2006 surgieron voces críticas que reprochaban el avance lento del W3C"

Antecedentes

Html es el lenguaje de marcación de hipertexto que se diseñó en 1989 para crear páginas web. Desde su creación ha ido evolucionando de acuerdo con diversas necesidades que se orientan sobre todo a mejorar el procesamiento de la información, y así fue como aparecieron varias revisiones (html 2, html 3.2, html 4, html 4.01) que ampliaron y depuraron este lenguaje.

Xhtml 1 se estableció en enero de 2000 como una reformulación de html con sintaxis xml. Desde entonces parecía evidente que los futuros desarrollos de html se harían bajo la bandera del xhtml, y así consta en numerosos manuales publicados a partir de esa fecha. Durante mucho tiempo xhtml se presentó como la evolución natural de un html que debería dejar de usarse.

En agosto de 2002 el W3C comenzó a definir xhtml 2, una actualización de xhtml 1 con varias mejoras semánticas y una lógica más robusta, pero sin compatibilidad regresiva. Era realmente un lenguaje nuevo y así lo confirmaba el hecho de que tenía un nuevo espacio de nombres. Pero a finales de 2009 el W3C detuvo el desarrollo de xhtml 2.

Espacio de nombres

Es un conjunto de nombres, todos únicos. El uso de espacios de nombres (*namespace*) es una recomendación del W3C para resolver la ambigüedad que se puede producir entre elementos y atributos que se llamen igual en vocabularios xml distintos.

Por ejemplo, el elemento *title* de un documento xhtml expresa el título del documento (dentro del elemento *head*), mientras que el elemento *title* de un canal rss se utiliza para representar el título del canal (dentro del elemento *channel*) y el título de cada uno de los artículos (dentro del elemento *item*).

Un grupo de diseñadores de *Opera*, y más tarde de *Apple* y *Mozilla*, estaban disconformes con el rumbo del W3C y en 2004 constituyeron el *Web Hypertext Application Technology Working Group (WhatWG)* al margen del W3C. Su propósito era crear una nueva versión del estándar desde un punto de vista eminentemente práctico, y no tan académico. Así, publicaron dos especificaciones, *Web Forms 2.0* y *Web Applications 1.0*, que más tarde se unirían para constituir un borrador de html5 (Lawson; Sharp, 2011).

A mediados de 2006 surgieron voces críticas que reprochaban el avance lento del W3C, la falta de resultados concretos y lo abstracto de sus propuestas (Castro, 2007). En efecto, 7 años después de que el W3C publicase las recomendaciones oficiales de html 4.01 y xhtml 1, aún no había ningún consenso salido del W3C sobre el camino que se debía seguir.

“Html era un lenguaje poco riguroso y como consecuencia los documentos resultaban un tanto caóticos”



Finalmente en 2007 el W3C reconoció el trabajo del *WhatWG* y lo tomó como base para su propia actividad con el fin de proponer una actualización

del estándar básico de la Web. A pesar de que el desarrollo de html estaba cerrado, el W3C se concentró en definir html5. Desde entonces ha habido cierta confusión y acalorados debates sobre cómo debería ser el lenguaje para el futuro de la Web.

Sintaxis y corrección de errores

Xhtml se creó para sustituir la sintaxis en la que se basa html: el sgml (*standard generalized markup language*), cuya finalidad principal era facilitar la creación manual de documentos. Html era un lenguaje poco riguroso y como consecuencia los documentos resultaban un tanto caóticos y no siempre cumplían con la sintaxis. Esto provocó que los navegadores se hicieran más complejos para poder asumir ambigüedades y equivocaciones, es decir, para poder representar los documentos aunque tuvieran errores.

Como las máquinas tienen ciertas dificultades para leer y manipular contenido etiquetado con la sintaxis de sgml, se propuso una sintaxis nueva basada en xml, un lenguaje derivado de sgml pero más comprensible para las máquinas. Así, al aplicar a html la sintaxis de xml, los ordenadores son capaces de manipular xhtml con facilidad y precisión, y en particular los dispositivos móviles pueden procesar la información de manera más eficiente.

“Html5 incluye elementos nuevos destinados a enriquecer la presentación de documentos”

La especificación de html5 no se adscribe a una sintaxis o a la otra, sino que admite ambas: html y xhtml. De esta manera, los creadores de contenido pueden escoger entre un enfoque práctico aunque poco riguroso (sintaxis html) y una visión académica y estricta (sintaxis xhtml). Con el tiempo, el W3C ha acabado aceptando que html y xhtml sean recomendaciones paralelas que pueden coexistir.

Html5 no sólo define cómo se deben analizar los documentos, sino también cómo se deben interpretar si no son válidos o si están mal formados. Actualmente los navegadores corrigen los errores de sintaxis de distinta manera, de modo en que a los fabricantes les resulta más práctico. Html5 trata de poner fin a esa necesidad de ingeniería inversa de los navegadores, que compiten por definir cómo se deben subsanar los errores (Andersson, 2007; Keith, 2010).

DOM

Una de las novedades principales de html5 es la inclusión del *document object model* (DOM) como fundamento del lenguaje. DOM describe la estructura de un documento de acuerdo con el paradigma de la orientación a objetos. En otras palabras, define el conjunto de entidades que están presentes en un documento html y las acciones que pueden realizarse sobre ellas. Hasta ahora el DOM siempre se había tratado de forma separada; cada navegador lo implantaba según la particular interpretación que su fabricante hacía de él. En cambio en html5 el DOM forma parte del estándar. Así se garantiza que los navegadores interpretarán adecuadamente la sintaxis de html y que al mismo tiempo implantarán las funciones del DOM que la sustentan.

Si el DOM está vinculado al lenguaje html, se evita tener que diseñar versiones distintas de una misma página para varios programas. El DOM se incluye mediante la extensión de sus API (interfaces de programación de aplicaciones), a las que añade funciones nuevas. Esta ampliación de las capacidades del DOM permite ejecutar funciones sofisticadas que hasta ahora requerían programas y componentes adicionales (Álvarez-García, 2010). Por ejemplo, Google prevé sustituir con mecanismos propios de html su extensión *Gears*, desarrollada para sincronizar contenidos para ser usados sin conexión.

Para describir la interfaz de los elementos del DOM con un lenguaje neutro, los autores de html5 han escogido el *interface definition language* (IDL). Este lenguaje tiene una sintaxis similar a C++, que proporciona métodos para definir conceptos asociados a la programación orientada a objetos: atributos, métodos, constantes, herencias, etc. Pero IDL presenta una desventaja importante: carece de mecanismos para especificar la jerarquía que debe haber entre elementos y las restricciones (obligatoriedad y valores posibles) que cabe aplicar a los atributos. No obstante, los autores de html5 han escogido este lenguaje, en lugar de DTD o xml schema, porque puede definir elementos y comportamientos sin depender de una sintaxis concreta, una tarea imposible para DTD y xml schema (Korostov; Paramzin, 2010).

Presentación y recogida de información

Html5 incluye elementos nuevos destinados a enriquecer la presentación de documentos. Son ejemplos de ello los elementos semánticos *article*, *header*, *hgroup*, *nav*, *section*, *aside* y *footer*. Con ellos se pretende evitar que los autores abusen del elemento *div* para delimitar partes de

un documento. Los blogs y los sitios de noticias han influido en gran medida en esta evolución (Schafer, 2010).

Por otra parte existe un medio para que el usuario pueda hacer llegar datos de entrada a un servidor: los formularios, que recogen información que después remiten a aplicaciones que se ejecutan en el servidor (CGI, API, JSP, *servlets* u otras interfaces). Html5 define más de una docena de nuevos controles (*email*, *range*, *date*, *time*, *placeholder*, *autofocus*, etc.) que actúan por fin sin necesidad de utilizar *JavaScript*, un lenguaje de programación que no todos los usuarios tienen habilitado (Pilgrim, 2010).

Html5 propone por tanto que sean los navegadores –y no los creadores de contenido– quienes faciliten la entrada y la validación de datos que tienen un patrón regular o están sometidos a restricciones. Así, el dolor de cabeza que supone para los diseñadores de páginas web verificar el formato de direcciones electrónicas, intervalos de valores, términos de búsqueda, colores, fechas y horas, entre otros tipos de datos, se descarga ahora sobre los navegadores.

“Sin que *Flash* vaya a desaparecer, seguramente html5 lo desplazará como contenedor de vídeo y audio”

Flash y contenido multimedia

Para incrustar contenido multimedia, html ya contaba con el elemento *object*, pero la nueva versión del estándar hace una propuesta más semántica. En el ámbito del multimedia, html5 incorpora directivas nuevas que actúan como contenedores de vídeo, gráficos vectoriales y audio. Se espera que con los elementos *video* y *audio* los navegadores tengan la capacidad de presentar el contenido de forma nativa, es decir, sin requerir componentes externos como *Flash*. Al fin y al cabo, una tarea sencilla como representar contenido multimedia no tiene por qué quedar en manos de un entorno cerrado y propietario.

Este avance ha suscitado el debate sobre si html5 acabaría con *Flash*. Es cierto que la propuesta de html5 puede sustituir funciones que hasta ahora sólo eran posibles con *Flash*. Para reproducir multimedia no hace falta ningún mecanismo complejo, por lo que es razonable pensar que html5 desplazará a *Flash* como contenedor de vídeo y audio, pero no hay argumentos sólidos para pensar que éste vaya a desaparecer (Allaire, 2010). Al fin y al cabo, *Flash* no sólo sirve para mostrar



contenido multimedia, sino que también tiene capacidad para resolver muchas otras tareas. El objetivo de la propuesta de

html5 es que la reproducción de vídeo y audio sea más eficiente, consuma menos recursos y se pueda gestionar con un código abierto y transparente, y sin necesidad de instalar componentes adicionales.

Todo apunta a que los navegadores incorporarán estas capacidades y que muchas páginas web se modificarán para aprovecharlas. Pero de momento, aunque incluidas en el estándar, esas funcionalidades se encuentran en la primera etapa de su adopción y no permiten controlar el contenido multimedia en la misma medida que *Flash*. De hecho, el popular depósito de vídeos *YouTube* ha declarado recientemente que html5 no cubre todas sus necesidades (**Harding**, 2010). Varios son los problemas: html5 no cuenta todavía con un formato de vídeo estándar –se disputan ese puesto *H.264*, *Ogg Theora* y *WebM*–, no sustenta el visionado a pantalla completa ni permite escoger la calidad de reproducción. Por lo tanto, aunque las novedades de html5 son un avance hacia los estándares abiertos, es razonable pensar que *Flash* continuará desempeñando un papel vital en la distribución de contenido multimedia.

Presente y futuro

Html5 se presenta como un marco estable para el diseño de páginas web. Ha sido definido para durar muchos años gracias a diversos mecanismos que permiten extender el lenguaje con facilidad. La filosofía de html5 aboga además por los estándares abiertos, que son fundamentales para impulsar la innovación e introducir en la sociedad los beneficios de las nuevas tecnologías. Bien sabemos que los estándares abiertos tienen la capacidad de mejorar nuestras vidas, y de ello son ejemplo las nuevas funciones multimedia de html5.

Debido al uso extendido de determinados complementos, como el citado *Flash*, las nuevas funciones multimedia de html5 no pueden aspirar a sustituirlos por completo –como ya se ha dicho–, pero facilitan otros métodos para presentar contenido multimedia. La intención es no depender tanto de los complementos externos y propietarios.

Con todo, el proyecto de html5 ha comenzado a ir un poco a la deriva. Algunos navegadores han comenzado a exhibir desarrollos propietarios

(**Powers**, 2010), lo cual amenaza con socavar la filosofía fundamental. Además, la aparente escisión entre el *W3C* y el *WhatWG* sugiere que html5 podría andar fuera del camino deseado, o podría emitir un estándar diferente del que la industria necesita (**Huggers**, 2010).

El *W3C* y los fabricantes de navegadores representan el futuro de la Web. Html5 debe cumplir con su promesa de favorecer un navegador estándar, abierto y único. Ésta debe ser la preocupación principal, más allá del dilatado debate, quizá excesivo, en torno a la reproducción de vídeo. Aún hay mucho por hacer en html5 antes de poder integrarse en todos los ámbitos de la Web. Hay tiempo suficiente para volver al buen camino.

El último borrador de la especificación html5 se puede consultar en:

<http://dev.w3.org/html5/spec>

Bibliografía

Allaire, Jeremy. "The future of Web content: html5, Flash and mobile apps". *TechCrunch*, 5 febrero 2010. <http://techcrunch.com/2010/02/05/the-future-of-web-content-html5-flash-mobile-apps>

Álvarez-García, Alonso. *Html5*. Madrid: Anaya Multimedia, 2010.

Andersson, David. "Html5, xhtml 2 and the future of the Web". *Digital web magazine*, 10 abril, 2007. http://digital-web.com/articles/html5_xhtml2_and_the_future_of_the_web

Castro, Elizabeth. *Html, xhtml y CSS*. Anaya Multimedia, 2007.

Harding, John. "Flash and the html5 <video> tag". *YouTube API Blog*, 29 junio 2010. <http://apiblog.youtube.com/2010/06/flash-and-html5-tag.html>

Huggers, Erik. "Html5, open standards, and the BBC". *BBC internet blog*, 13 agosto 2010. http://www.bbc.co.uk/blogs/bbcinternet/2010/08/html5_open_standards_and_the_b.html

Keith, Jeremy. *Html5 for web designers*. Nueva York: A Book Apart, 2010.

Ejemplo de web construida en html5

Se trata de un videoclip experimental interactivo.

Se recomienda utilizar *Google Chrome*:

<http://thewildernessdowntown.com>

Enviado a ThinkEPI por **Antón Suárez-Beltrán**

asuarzbeltran@yahoo.es

Korostov, Alexander V.; Paramzin, Dmitry J. "The arrival of html5". *Dr. Dobbs*, 13 agosto 2010.
<http://www.drdobbs.com/web-development/226700204>

Lawson, Bruce; Sharp, Remy. *Introducing Html5*. Berkeley: New Riders: 2011.

Pilgrim, Mark. *Html5: up and running*. Cambridge

[etc.]: O'Reilly Media, 2010.

Powers, Shelley. "Apple, Opera, and Mozilla: why are you working against open standards?". *Burningbird*, 10 de junio, 2010.
<http://burningbird.net/node/108>

Schafer, Steven M. *Html, xhtml y CSS*. Madrid: Anaya Multimedia, 2010.

El profesional de la información
colaborará con la **Conferencia VISIO 2011**,
publicando una selección de sus mejores comunicaciones
(que serán inéditas) en nuestro número sobre inteligencia
competitiva de septiembre 2011.



El 2 y 3 de junio de 2011 se celebrará en la Alhóndiga de Bilbao, la 4ª edición de la *Conferencia Internacional VISIO 2011 - Vigilancia e Inteligencia Sistémica para la Innovación en las Organizaciones*.

<http://www.conferencia-visio.com/2011/visio2011.php?id=1&pag=presentacion>



Ésta tiene como objetivo reunir a profesionales que contribuyan a aumentar la calidad y éxito de los expertos en Vigilancia e Inteligencia, nacionales y/o extranjeros.

