HTML5: A New Standard for the Web

Matthew B. Hoy

Routledge
Taylor & Francis Group

NAVIGATE THE NET

Donna Berryman and Matthew B. Hoy, Column Editors

HTML5: A New Standard for the Web

MATTHEW B. HOY

Luther Midelfort-Mayo Health System, Eau Claire, Wisconsin

HTML5 is the newest revision of the HTML standard developed by the World Wide Web Consortium (W3C). This new standard adds several exciting news features and capabilities to HTML. This article will briefly discuss the history of HTML standards, explore what changes are in the new HTML5 standard, and what implications it has for information professionals. A list of HTML5 resources and examples will also be provided.

KEYWORDS   HTML, HTML5, standards

BRIEF HISTORY OF HTML

HTML, or Hyper-Text Markup Language, is what creates the basic structure of the World Wide Web. HTML "consists of encoded markers called tags that surround and differentiate bits of text, indicating the function and purpose of the text those tags 'mark up.'"[1] Some tags specify how information is displayed within the document; others create links between documents and other content on the Internet. For example, the <img> tag inserts an image into the document; the <a> or anchor tag creates a link to another document.

HTML was first developed in 1989 by Tim Berners-Lee, the widely acknowledged inventor of the web. He created HTML to allow researchers

at geographically distant institutions to collaborate and share large amounts of data electronically. Berners-Lee based his new language on an already existing one called Standard Generalized Mark-up Language (SGML). He borrowed many of the key layout concepts and tags in HTML directly from SGML, including titles, paragraphs, headings, ordered lists, and many others. His major breakthrough was the development of hypertext links in HTML; "the idea of using the anchor element with the HREF attribute was purely Tim's invention, as was the now-famous "www.name.name format for addressing machines on the web."[2] These hypertext linkages are what allows HTML to quickly link to documents stored anywhere on the Internet.

The early 1990s saw a number of competing web browsers being developed, and many of them expanded and extended the capabilities of Berners-Lee's initial HTML design. This led to problems; pages designed for one browser displayed poorly or not at all in other browsers. In response to this splintering of the initial HTML standard, Tim Berners-Lee formed the World Wide Web Consortium (W3C) in 1994. The W3C's stated mission is "to lead the World Wide Web to its full potential by developing protocols and guidelines that ensure the long-term growth of the Web."[3] To achieve that mission, the W3C develops open standards for many web technologies, including HTML. Lee and the other members of the W3C believe that by maintaining open standards, they will keep the web free and "available to all people, whatever their hardware, software, network infrastructure, native language, culture, geographical location, or physical or mental ability."[3]

There were several early versions of HTML; Berners-Lee's original design was called simply HTML, HTML 2.0, HTML+, and HTML 3.0. While these versions were widely adopted and used throughout the industry, they were never officially codified as standards. The first true standard for HTML was published by the W3C in 1998. This standard, called HTML 4, added many new features including style sheets, frames, embedded objects, complex forms, and page accessibility tools.[4] Newer standards, based on Extensible Markup Language (XML), were developed in 2000 and revised periodically over the last ten years. These included XHTML 1.0, 1.1, Basic, and 2.0.

During the relatively short time that there have been standards for HTML, there have been fundamental changes in the way the web is used. When the first standard was released, little thought was given to concepts such as multimedia support and mobile devices; the web was primarily a source for static documents and accessed exclusively by computers. Now, users want dynamic pages that behave more like applications than documents. They want to use the web on different devices and platforms. They expect high quality audio and video and want to access content from anywhere. Web designers use a mish-mash of plug-ins and extensions to add these features, and users are required to install extra software and troubleshoot compatibility issues. To address these and other problems, the W3C is developing a new version of the HTML standard: HTML5. The following

section will describe some of the changes and new features being added in HTML5.

NEW FEATURES IN HTML5

The W3C is designing the HTML5 standard with several goals in mind: replacing proprietary multimedia plug-ins with open standards, allowing web applications to behave more like native applications, adding features for location-based services, and making changes to the syntax that separate content from presentation.

The most talked about new feature in HTML5 is native support for audio and video playback in the browser. Current browsers rely on plug-ins such as Adobe Flash, Apple Quicktime, and Microsoft Silverlight to play multimedia content. For desktop users, these plug-ins work well and are relatively transparent to use; for the growing number of users who want to view video on their iPhones, iPads, or Blackberries, there are no plug-ins available. HTML5 will solve the plug-in problem by mandating built-in support for video in browsers and allowing designers to insert <video> and <audio> tags into their pages. These tags will contain information the browser can use to interpret and render the multimedia content. Although a final encoding standard for video has not been chosen, the most likely candidates are currently Ogg Theora and H.264, both widely used open video standards. Apple's Safari browser group and Microsoft's Internet Explorer team are currently planning to support H.264, while Mozilla Firefox and the Opera browser are backing Ogg Theora. Google's Chrome browser will support both.[5]

Another much-anticipated new feature of the HTML5 standard is the ability to build web-apps that behave more like local applications. Features such as drag-and-drop file management and local storage will allow developers to build web-based applications that integrate seamlessly with a user's computer. Users will be able to upload and download files by simply dragging them to and from the browser window. HTML5 will also have support for client-side SQL databases and offline caching. This is particularly important for mobile users, since most have limited connectivity and bandwidth on their devices. An example of this "locally cached application" concept would be a calendar suite that allows users to download their current data, make adjustments on their mobile device, and then upload the next time there is a connection available. Applications can store graphics and code on the local device, ensuring quick load times and minimal network traffic.

The Canvas element is another new feature of HTML5. Developers will be able to combine vector graphics, images, audio, and video within a predefined space on the page. These elements can be programmed via JavaScript, and users will be able to interact with them via mouse and keyboard. Rendering will be handled on the client side, ensuring smoother processing and

fewer network bottlenecks. The canvas element can contain things as simple as geometric shapes or lines, and things as complex as games and layered models that can be manipulated by the user.[6]

HTML5's Geolocation features will be of particular interest to mobile developers and users. Geolocation has been defined as "the art of figuring out where you are in the world and (optionally) sharing that information with people you trust."[7] By including tools for geolocation in the browser, HTML5 will make it easy to create applications that know where users are, what is close by, and how they can get where they want to go. Future applications could include maps and directions (where am I and how can I get somewhere else?), shopping and entertainment (where is the closest restaurant and is it open this late?), and social networking tools (who is close by and what do we have in common?). Although there are legitimate privacy concerns with these features, it should be noted that the current design specification requires that users explicitly choose to allow geolocation applications to know where they are.

The other changes coming in HTML5 deal with refining the syntax of the language. Earlier versions of HTML combined presentational markup and semantic markup in one document. HTML5 deprecates presentational markup such as font color and size tags in favor of style sheets. Although style sheets have been supported since HTML4, this new standard really pushes developers to separate presentation from content. The move makes sense because web pages that use style sheets are more consistent and much easier to maintain. HTML5 also adds some new semantic tags like <article>, <aside>, and <nav> that will better define blocks of content within pages.

As with all previous updates to the HTML standard, HTML5 browsers will be backward compatible. Older pages will continue to display in new browsers, and there will be no need to rewrite any existing documents.

HTML5 ISSUES

Although there is a great deal of enthusiasm about the new features of HTML5, there are several hurdles to be overcome. The largest hurdle is the glacial pace of standards development and implementation. The W3C has been working on the HTML5 standard for several years. It is not due to become a recommendation until 2012 and will not be a finalized standard until 2022 at the earliest. Although some of the features have already been implemented in several browsers, the long delay before a final standard is reached may cause browser developers to abandon some features or extend others, creating more consistency problems.

Even if there is a completed standard and standards-compliant browsers are available, it will take years for the new standard to be rolled out. A widely accepted statistic is that new features must be supported by 80% of deployed

browsers in order to be considered reliable. Considering how rarely most users upgrade their browser, it could take several years to reach a critical mass of deployed browsers. It is estimated that the earliest HTML5 compliant browsers will reach this threshold is 2013.[6]

HTML5's blurring of the line between desktop and web applications creates another issue. By allowing data and code from web apps to be stored and executed on the local device, it opens a number of security holes. Previous versions of HTML only allowed small text files called cookies to be stored locally. At this early stage in the process, there are few verified exploits, but users and developers will need to be aware of the security implications of HTML5 and watch for future problems.[8]

CONCLUSION

HTML5 offers several new and interesting features, although it may be several years before these features are universally supported. Native audio and video, a programmable canvas area within pages, and extensive geolocation features will undoubtedly lead to the development of exciting new web applications and tools. The addition of <article> and <aside> tags will benefit researchers by providing more context within pages and allow for easier programmatic gathering of citation information. The enhanced multimedia and scripting support will allow designers to create a richer web experience, especially for mobile users. Readers who develop web pages should familiarize themselves with the new standards and consider how they might include these features in their future projects. Everyone else should download an HTML5-compliant browser and get ready for the next generation of websites.

HTML5 RESOURCES

The following resources offer a more detailed look at the HTML5 standard:

- Draft of the HTML5 Standard at the World Wide Web Consortium: <http://dev.w3.org/html5/spec/Overview.html>
- *Dive into HTML5* by Mark Pilgrim, a free e-book explaining the core concepts of HTML5: <http://diveintohtml5.org/>
- HTML5 Demos, a collection of demos that highlights new features and gives detailed information about which browsers currently support HTML5: http://html5demos.com/

REFERENCES

1. Schulz, D., and Craig, C. "*Beginning HTML with CSS and XHTML: Modern Guide and Reference.*" Berkeley: Apress, 2007.

2. Ragget, D.; Lam, J.; Alexander, I.; and Kmiec, M. "*Raggett on HTML 4.*" Boston: Addison Wesley, 1998.

3. W3C. "W3C Mission." Available: <http://www.w3.org/consortium/mission>. Accessed: September 3, 2010.

4. Raggett, D.; Le Hors, A.; and Jacobs, I. "HTML 4.01 Specification." 1999. Available: <http://www.w3.org/TR/html4/>. Accessed: September 10, 2010.

5. Ozer, J. "Here Comes HTML5 – Should We Care?" *EventDV* 23 (May 2010): 48.

6. Vaughan-Nichols, S. J. "Will HTML5 Restandardize the Web?" *Computer* 43, no. 4 (2010): 13–15.

7. Pilgrim, M. "You Are Here (And So Is Everybody Else) – Dive Into HTML5." 2010. Available: <http://diveintohtml5.org/geolocation.html>. Accessed: September 12, 2010.

8. Mansfield-Devine, S. "Divide and Conquer: The Threats Posed by Hybrid Apps and HTML5." *Network Security* 2010, no. 3 (March 2010): 4–6.