

Will HTML 5 Restandardize the Web?

➔ Steven J. Vaughan-Nichols



The World Wide Web Consortium is developing HTML 5 as a standard that provides Web users and developers with enhanced functionality without using the proprietary technologies that have become popular in recent years.

In theory, the Web is a resource that is widely and uniformly usable across platforms. As such, many of the Web's key technologies and architectural elements are open and platform-independent.

However, some vendors have developed their own technologies that provide more functionality than Web standards—such as the ability to build rich Internet applications.

Adobe System's Flash, Apple's QuickTime, and Microsoft's Silverlight are examples of such proprietary formats.

In addition, Google's Gears and Oracle's JavaFX—which the company acquired along with Sun Microsystems—have technologies that enable creation of offline and client-side Web applications.

Although these approaches provide additional capabilities, they have also reduced the Web's openness and platform independence, and tend to lock in users to specific technologies and vendors.

In response, the World Wide Web Consortium (W3C) is developing HTML 5 as a single standard that provides Web users and developers with

enhanced functionality without using proprietary technologies.

Indeed, pointed out Google researcher Ian Hickson, one of the W3C's HTML 5 editors, "One of our goals is to move the Web away from proprietary technologies."

The as-yet-unapproved standard takes HTML from simply describing the basics of a text-based Web to creating and presenting animations, audio, mathematical equations, typefaces, and video, as well as providing offline functionality. It also enables geolocation, a rich text-editing model, and local storage in client-side databases.

The Web isn't just about reading the text on the page and clicking on the links anymore, noted Bruce Lawson, standards evangelist at browser developer Opera Software.

Added W3C director Tim Berners-Lee, "HTML 5 is still a markup language for webpages, but the really big shift that's happening here—and, you could argue, what's actually driving the fancy features—is the shift to the Web [supporting applications]."

"HTML 5 tries to bring HTML into the world of application development," explained Microsoft senior principal architect Vlad Vinogradsky.

"Microsoft is investing heavily in the W3C HTML 5 effort, working with our competitors and the Web community at large. We want to implement ratified, thoroughly tested, and stable standards that can help Web interoperability," said Paul Cotton, cochair of the W3C HTML Working Group and Microsoft's group manager for Web services standards and partners in the company's Interoperability Strategy Team.

At the same time though, Web companies say their proprietary technologies are already up and running, unlike HTML 5.

Adobe vice president of developer tools Dave Story said, "The HTML 5 timeline states that it will be at least a decade before the evolving efforts are finalized, and it remains to be seen what parts will be implemented consistently across all browsers."

In fact, while HTML 5 recently became a working draft, it's not expected to become even a W3C candidate recommendation until 2012 or a final W3C standard until 2022.

Nonetheless, some browser designers, Web authors, and websites—such as YouTube—are already adopting HTML 5 elements. For more

➔ PLAYING WITH HTML 5

Several websites offer a taste of what HTML 5 will bring. Some of the applications work with only certain browsers.

YouTube's beta HTML 5 video project (www.youtube.com/html5) works with the standard's video tag. Browsers must support the video tag and have a player that uses the H.264 codec. Otherwise, YouTube will use Flash to play video.

Mozilla Labs' BeSpin (<https://bespin.mozillalabs.com>) is an experimental programmer's editor that uses a variety of HTML 5 elements.

FreeCiv.net (www.freeciv.net) is an online game by the FreeCiv.net open source project that supports HTML 5's Canvas element. HTML 5-compatible browsers display map changes faster than those that aren't compatible.

Google Wave (<https://wave.google.com/wave>), a cross between social networking and groupware, uses several HTML 5 elements.

Merge Web Design's HTML 5 Geolocation (<http://merged.ca/iphone/html5-geolocation>) is, as the name indicates, a demo of HTML 5-based geolocation.

Sticky Notes (<http://webkit.org/demos/sticky-notes/index.html>) is the WebKit Open Source Project's demo of HTML 5's client-side database storage API. WebKit is an open source Web browser engine now used by, for example, Apple's Safari browser.

information, see the "Playing with HTML 5" sidebar.

BACKGROUNDER

HTML is the predominant markup language for webpages. It uses tags to create structured documents via semantics for text—such as headings, paragraphs, and lists—as well as for links and other elements. HTML also lets authors embed images and objects in pages and can create interactive forms.

HTML, which stemmed from the mid-1980s Standardized General Markup Language, first appeared as about a dozen tags in 1991.

The Internet Engineering Task Force began the first organized effort to standardize HTML in 1995 with HTML 2.0.

The IETF's efforts to maintain HTML stalled, so the W3C took over standardization.

In 1996, the W3C released HTML 3.2, which removed the various proprietary elements introduced over time by Microsoft and Netscape.

HTML 4.0, still not a final standard, followed in 1998. The approach provides mechanisms for style sheets, scripting, embedded objects, richer tables, enhanced forms, and improved accessibility for people with disabilities.

However, HTML was still primarily focused on delivering text, not multimedia or client-based applications.

Because of this, proprietary technologies such as Apple's QuickTime and Microsoft's multimedia players, both first released in 1991; and Adobe Flash, which debuted in 1996, have been used for video.

Technologies such as Google Gears and Oracle's JavaFX, both first released in 2007, make creating Web-based desktop-style applications easier for developers.

HTML 5

W3C is designing HTML 5 to create a standard with a feature set that handles all the jobs that the proprietary technologies currently perform, said specification editor Michael Smith, the consortium's special-missions-subsection junior interim floor manager.

In addition, HTML 5 will support newer mobile technologies such as geolocation and location-based services (LBS), as well as newer open formats such as scalable vector graphics. SVG, an open XML-based file format, produces compact and high-quality graphics.

Developers would thus be able to develop rich webpages and Web-based applications without needing to

master or license multiple proprietary technologies.

And browsers would be able to do more without plug-ins.

Canvas. One of HTML 5's key new features is Canvas, which lets developers create and incorporate graphics, video, and animations, usually via JavaScript, on webpages.

HTML Canvas 2D Context is an Apple-originated technology for rendering 2D graphics and animations on the client rather than on Web servers.

By rendering graphics locally, the bottlenecks of server and bandwidth restrictions are avoided. This makes graphics-heavy pages render faster.

Video tags. HTML 5's codec-neutral video tags provide a way to include nonproprietary video formats, such as Ogg Theora and H.264, in a page.

The tag and underlying code tell the browser that the associated information is to be handled as an HTML 5-compatible video stream.

They would also let users view video embedded on a webpage without a specific video player.

Location-based services. A location API offers support for mobile browsers and LBS applications by enabling interaction with, for example, GPS technology and data.

Working offline. AppCache lets online applications store data and programming code locally so that Web-based programs can work as desktop applications, even without an Internet connection.

HTML 5 has several other features that address building Web applications that work offline. These include support for a client-side SQL database and for offline application and data caching.

Web applications thus can have their code, graphics, and data stored locally.

Web Workers. The Web Workers element runs scripts in the background that can't be interrupted by other scripts or user interactions. This speeds up background tasks.

Syntax and semantics. HTML 5 makes changes to the syntax and the semantics of the language's elements and attributes.

For example, as Figure 1 shows, HTML can be written in two syntaxes: HTML and XML.

Using XML will enable more complex webpages that will run faster on Web browsers.

XML requires a stricter, more accurate grammar than HTML and thus necessitates less work by the local computer to run quickly and correctly.

However, XML pages require more work by the developer to achieve the higher accuracy level.

HTML+RDFa. The W3C recently began dividing HTML 5 into subsections for easier development.

For example, HTML + RDFa (Resource Description Framework in attributes) provides a way to embed resource description frames in webpages.

RDF is a method for conceptually describing or modeling information implemented in Web resources.

This would let developers incorporate machine-readable data into webpages, which would enable faster page rendering.

TRIALS AND TRIBULATIONS

Both nailing down details and the standardization process itself have been problematic for HTML 5.

For example, the open source community supports making the open Ogg Theora format the default video codec for HTML 5's video tag.

Adobe, Google, and Microsoft opposed this, expressing video-quality, patent-related, and other concerns.

According to Google's Hickson, "There is no suitable codec that all vendors are willing to implement and ship."

The W3C thus decided not to select one as a default.

What's really slowing HTML 5 is that the standardization effort "combines the worst elements of the

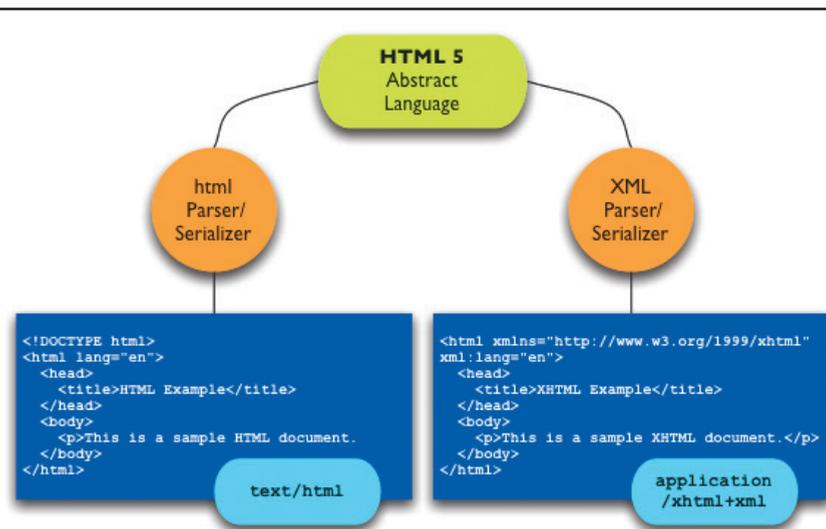


Figure 1. HTML 5 lets HTML be written in two syntaxes: HTML and XML. XML's stricter, more accurate grammar requires more work by the developer but less work by the local computer to run quickly and correctly. The browser first parses out the various terms in HTML and XML and then sends them serially in an efficient and logical manner for rendering.

IETF process and the W3C process," said Adobe principal scientist Larry Masinter, a member of the W3C's HTML Working Group.

"[As is the case with] the IETF [process], there is the chaos of an open mailing list, wide-ranging comments, and free participation but without the 'adult supervision' that the IETF supplies in the form of the Internet Engineering Steering Group and area directors," he explained.

From the W3C, he added, there's too much bureaucracy and the voting members don't have enough responsibility and accountability for the final results.

Another potential implementation concern is that some companies are already extending HTML 5 in ways that the final standard is unlikely to support.

HTML 5 will blur the line between desktop and online applications and thus create an opportunity for malware writers, according to Dmitri Alperovitch, vice president of threat research with security vendor McAfee.

By letting Web applications run on local systems, HTML 5 would allow Web-based malware to do the same.

Timely user acceptance may be an issue for HTML 5.

Trevor Lohrbeer, CEO of Lab Escape, a data-visualization-tool vendor, said sweeping Internet technologies must be supported by at least 80 percent of browsers before being considered reliable enough for deployment.

He estimated HTML 5 won't achieve this level of support until 2013 at the earliest.

However, Jay Baker, director of architecture at Viewlocity, a supply-chain-management software vendor, said his company plans to adopt HTML 5 fairly aggressively.

"Although there is a learning curve associated with HTML 5," he explained, "many of the new features make today's common hacks and workarounds obsolete." **■**

Steven J. Vaughan-Nichols is a freelance technology writer based in Mills River, North Carolina. Contact him at sjvn@vna1.com.

Editor: Lee Garber, [Computer](mailto:l.garber@computer.org);