

[Webware](#)

July 8, 2009 4:00 AM PDT

An epitaph for the Web standard, XHTML 2

by [Stephen Shankland](#)

XHTML 2, we hardly knew you.

154
diggs

[digg it](#)

XHTML 2, a technology intended to build a more powerful Web from the ground up, met a quiet end last week, spotlighting the difficulties of standardization in a fast-moving Internet. Introduced in 2002, XHTML 2 was a centerpiece of standards work at the [World Wide Web Consortium \(W3C\)](#).

But incompatibility with the existing Web and a direction at odds with Web developers' desires doomed it to a slow demise. On Thursday, after a long reconciliation with browser makers who'd struck off in a different direction, the [W3C announced](#) that it will wind down development of XHTML 2 this year.

Instead, the group will channel those resources into standardizing what the browser makers have been toiling on all these years: HTML 5, a sprawling collection of new features to improve the present



Ultimately, Web browser makers had the upper hand in charting the Web's future.

(Credit: Stephen Shankland/CNET)

Hypertext Markup Language. Although elements of XHTML 2 will live on in HTML 5, overall, the browser makers prevailed.

"XHTML 2 was a beautiful specification of philosophical purity that had absolutely no resemblance to the real world," said Bruce Lawson, HTML 5 evangelist for browser maker Opera.

So what went wrong? In short, the Web has many masters, but the ones with final say over its nature are those who build it page by page, not the standards group trying to create a new foundation.

XHTML 2 was designed to reform the Web as a medium for publishing documents, but the developers--and the browser makers who listened closely to those developers--instead wanted a platform for interactive applications. And while that direction prevailed, its incarnation in HTML 5 faces its own set of challenges now.

The consensus for HTML 5 support has been building for years, and the W3C already had been increasing its involvement in its standardization well before it decided to put an end to much of the competing XHTML 2 standard. Although the HTML-XHTML split has been fractious at times, there's inescapable tension between standards groups trying to chart the future and vendors whose products relate to those standards.

"I will not say it's been the smoothest way of doing things, but it's not an unnatural way for things to proceed," said Mike Smith, leader of HTML work at W3C, speaking of the reconciliation process that rejuvenated the W3C's HTML work. "Vendors are the ones who drive innovation on the Web for the most part."

Why XHTML?

So if it's so clear today that HTML 5 is the way to go, why was so much energy, time, and research invested in XHTML 2? It was an attempt start afresh without HTML's shortcomings.

The X in XHTML stands for XML, which in turn stands for Extensible Markup Language. XML is a broad technology that uses a strict set of tags to label different types of content in a document, and XHTML was engineered specifically for the Web. XHTML brought rigor to the loosey-goosey and slap-dash world of HTML, and it would have permitted developers to employ a broader range of computing engines called parsers to digest and process the XML, Smith said.

XHTML "was a cleaner and better-architected version of HTML," Smith said. And in its earlier years, it had support. "At the time when XHTML 2 was first conceived and specified in the early drafts, most everybody thought it was a good idea. A lot of people in hindsight want to look back at it now and make the claim that they knew it wasn't going to have success," Smith said.

One example of its utility is the tight coupling of textual information with a graphs encoded with the [SVG, or Scalable Vector Graphics](#) format, Smith said. Another advantage was better browsing with the limited abilities of

XHTML™ 2.0

W3C Working Draft 26 July 2006

XHTML 2.0 made it to working draft stage, but only parts of the specification will live on in HTML

mobile phones.

5.

One of the big problems with XHTML 2 was that it wasn't backwards compatible, though. Not only could it not be used to display existing Web pages, but Web browsers had to be expanded with an entirely new engine for handling the XML. Notably, Microsoft's Internet Explorer, the dominant browser by far, couldn't handle XHTML on its own.

Another problem was that there was plenty of demand for improvements to HTML, which W3C had declared finished with version 4.01 in 1999.

"People were so focused on XHTML 2 that they were substantially less interested in modifying the application model and introducing new features to HTML that developers were clamoring for," said Arun Ranganathan, standards evangelist for Mozilla, the organization behind the [Firefox](#) browser. "We felt the standards going on at the time...were disconnected from a large majority of developers.

Microsoft agrees with its browser rival.

"We've never heard a strong request from our developer audience and customers for XHTML 2," said Amy Barzdukas, general manager for IE.

Enter WHATWG

One crucial moment came five years ago when Opera and Mozilla representatives showed the W3C an idea called WebForms for improving HTML. "We jointly presented this paper to W3C, who rejected it," Lawson said.

Mozilla's Brendan Eich and Opera's Ian Hickson were displeased with how things went. "The best way to help the Web is to incrementally improve the existing web standards," concluded Eich, founder of the JavaScript Web programming language, after the meeting in a [blog post](#).

Eich also announced there an Opera and Mozilla plan to take that evolutionary route. They launched an open e-mail list called [WHATWG](#), short for Web Hypertext Application Technology Working Group. Apple, which offers its own [Safari](#) browser, soon began participating, too.

"It became a de facto standards organization without the formality of W3C. It's where we went to figure out what the future of the Web was," Ranganathan said.

Eventually, the Web-application direction won over the W3C. "Some things are clearer with hindsight of several years. It is necessary to evolve HTML incrementally," said [Web](#)

[founder and W3C Director Tim Berners-Lee](#) said in 2006.

But Berners-Lee at the time also maintained the commitment to the "well-formed," more rigorous XML-based future: "It is important to maintain HTML incrementally, as well as continuing a transition to well-formed world, and developing more power in that world."



In practice, the W3C world and WHATWG world involve many of the same people. That probably eased the reconciliation to the current state, where WHATWG and W3C operate simultaneously, the first more informal and the second with more careful handling of intellectual property concerns.

Ultimately, HTML carried the day. What began with interest in more sophisticated Web sites such as eBay blossomed with the arrival of Ajax, which used JavaScript to build more sophisticated Web-based applications. And Web applications weren't just theoretical ideas.

"When Gmail and Google Maps and Ajax came along, it became really clear we needed a new set of technologies that made it easier to make those kinds of applications," Smith said.

The transition culminated with W3C's bare-bones news last week: "Today the director announces that when the XHTML 2 Working Group charter expires as scheduled at the end of 2009, the charter will not be renewed. By doing so, and by increasing resources in the HTML Working Group, W3C hopes to accelerate the progress of HTML 5 and clarify W3C's position regarding the future of HTML."

Some features of XHTML 2 will be built into HTML 5, so the XHTML 2 work won't have been for naught, assuming a critical mass of browser makers do in fact include the necessary XML parser along the HTML parser.

HTML 5: no walk in the park

Though the W3C-WHATWG dust has mostly settled, the standard is far from finished, and indeed looks a long way off.

The present approach involves a give and take between [browser makers trying out new features](#) and the standards group codifying them. Features can't make it to the ultimate W3C state, "final recommendation," until at least two browsers support the feature compatibly, Smith said.

In practice, that means adventurous Web developers who choose to support the new

technologies in effect are blessing them even though the technology might well change.

HTML 5 elements came from all over. Canvas, which involves two-dimensional graphics, began at Apple's Safari and now has won over Opera, Firefox, and Google's Chrome. ContentEditable, which lets Web pages be edited in place, came from Microsoft. Google now is working on a faster communication feature called Web Sockets. Programmers for WebKit, the open-source project underlying Safari, are developing DataGrid, which brings spreadsheet-like tables with sorting and editing to Web pages.

"The speed of the web is continuing to pick up in general," Barzdukas said. HTML 5 feature support figures prominently in the [browser sales pitches from Google](#) and from [Mozilla, with its "upgrade the Web" tag line for Firefox 3.5](#).

Actual standardization, though, remains distant. Mozilla's Ranganathan hopes for drafts of some HTML 5 elements this year and a draft of the full specification in 2010.

The HTML 5 built-in video situation is illustrative. Hickson, the HTML 5 editor and now Google employee, posted a [lament about HTML 5 video](#) last week because browser makers don't agree on whether to support the patent-free Ogg Theora format, preferred by Opera and Mozilla, or the commercially popular H.264 format, preferred by Google and Apple. The upshot for now: HTML 5 is trying to standardize video but doesn't specify which format to be used.

That pace of HTML 5 standardization important, given the importance Microsoft places on supporting actual standards and the company's commanding market share.

"The support of ratified standards (that Web developers) can use is something that we are extremely supportive of," Barzdukas said. "In some cases, it can be premature to start claiming support for standards that are not yet in fact standards."



Stephen Shankland writes about a wide range of technology and products, but has a particular focus on browsers and digital photography. He joined CNET News in 1998 and since then also has covered Google, Yahoo, servers, supercomputing, Linux and open-source software, and science. [E-mail Stephen](#), or follow him on Twitter at <http://www.twitter.com/stshank>.