

What Is XHTML, And Why Should I Care?

In prior posts in this series ([1](#), [2](#), [3](#), and [4](#)), I described with some glee my rash abandonment of common blogging software in favor of writing my own. For my purposes there seemed to be some key disadvantages^[1] to these popular packages, including an overemphasis on the calendar (I just saw the definition of a blog at the South by Southwest Interactive Festival—"a page with dated entries"—which, to paraphrase Woody Allen, is like calling *War and Peace* "a book about Russia"), a sameness to their designs, and comments that are rarely helpful and often filled with spam. But one of the greatest advantages of recent blog software packages is that they generally write standards-compliant code. More specifically, blog software like WordPress automatically produces XHTML. Some of you might be asking, what is XHTML, and who cares? And why would I want to spend a great deal of effort ensuring that this blog complied strictly with this language?

The large digital library contingent that reads this blog could probably enumerate many reasons why XHTML compliance is important, but I had two reasons in mind when I started this blog. (Actually, I had a third, more secretive reason that I'll mention first: Roy Rosenzweig and I argue in our book *Digital History*^[2] that XHTML will likely be critical for digital humanists to adhere to in the future—don't want to be accused of being a hypocrite.) For those for whom web acronyms are Greek, XHTML is a sibling of XML, a more rigorously structured and flexible language than the HTML that underlies most of the web. XHTML is better prepared than HTML to be platform-independent; because it separates formatting from content, XHTML (like XML) can be reconfigured easily for very different environments (using, e.g., different style sheets). HTML, with formatting and content inextricably combined, for the most part assumes that you are using a computer screen and a web browser. Theoretically XHTML can be dynamically and instantaneously recast to work on many different devices (including a

personal computer). This flexibility is becoming an increasingly important feature as people view websites on a variety of platforms (not just a normal computer screen, e.g., but cell phones or audio browsers for the blind). Indeed, according to the server logs for this blog, 1.6% of visitors are using a smart phone, PDA, or other means to read this blog, a number that will surely grow. In short, XHTML seems better prepared than regular HTML to withstand the technological changes of the coming years, and theoretically should be more easily preserved than older methods of displaying information on the web. For these and other reasons [a 2001 report the Smithsonian commissioned^{\[3\]}](#) recommended the institution move to XHTML from HTML.

Of course, with standards compliance comes extra work. (And extra cost. Just ask webmasters at government agencies trying to make their websites comply with Section 508, the mandatory accessibility rules for federal information resources.) Aside from a brief flirtation with the what-you-see-is-what-you-get, write-the-HTML-for-you program Dreamweaver in the late 1990s, I've been composing web pages using a text editor (the superb [BBEdit^{\[4\]}](#)) for over ten years, so my hands are used to typing certain codes in HTML, in the same way you get used to a QWERTY keyboard. XHTML is not that dissimilar from HTML, but it still has enough differences to make life difficult for those used to HTML. You have to remember to close every tag; some attributes related to formatting are in strange new locations. One small example of the minor infractions I frequently trip up on writing XHTML: the oft-used break tag to add a line to a web page must "close itself" by adding a slash before the end bracket (not `
`, but `
`). But I figured doing this blog would give me a good incentive to start writing everything in strict XHTML.

Yeah, right. I clearly haven't been paying enough attention to detail. The page you're reading likely still has dozens of little coding errors that make it fail strict compliance with the World Wide Web Consortium's XHTML standard. (If you would like a humbling experience that brings to mind receiving a pop quiz back from your third-grade teacher with

lots of red ink on it, try the [W3C's XHTML Validator](#)^[5].) I haven't had enough time to go back and correct all of those little missing slashes and quotation marks. WordPress users out there can now begin their snickering; their blog software does such mundane things for them, and many proudly (and annoyingly) display little "XHTML 1.0 compliant" badges on their sites. Go ahead, rub it in.

After I realized that it would take serious effort to bring my code up to code, so to speak, I sat back and did the only thing I could do: rationalize. I didn't *really* need strict XHTML compliance because through some design slight-of-hand I had already been able to make this blog load well on a wide range of devices. I learned from other blog software that if you put the navigation on the right rather than the more common left you see on most websites, the body of each post shows up first on a PDA or smart phone. It also means that blind visitors don't have to suffer through a long list of your other posts before getting to the article they want to read.

As far as XHTML is concerned, I'll be brushing up on that this summer. Unless I move this blog to WordPress by then.

Part 6: One Year Later^[6]

This entry was posted on Thursday, January 5th, 2006 at 10:13 pm and is filed under [Accessibility](#)^[7], [Blogs](#)^[8], [HTML](#)^[9], [Programming](#)^[10], [Software](#)^[11], [Standards](#)^[12], [Web](#)^[13]. You can follow any responses to this entry through the [RSS 2.0](#)^[14] feed. You can [leave a response](#)^[15], or [trackback](#)^[16] from your own site.

References

1. [^ seemed to be some key disadvantages](#) (www.dancohen.org)
2. [^ Digital History](#) (www.dancohen.org)
3. [^ a 2001 report the Smithsonian commissioned](#) (www.si.edu)
4. [^ BBEdit](#) (www.barebones.com)

5. [^ W3C's XHTML Validator \(validator.w3.org\)](#)
6. [^ Part 6: One Year Later \(www.dancohen.org\)](#)
7. [^ View all posts in Accessibility \(www.dancohen.org\)](#)
8. [^ View all posts in Blogs \(www.dancohen.org\)](#)
9. [^ View all posts in HTML \(www.dancohen.org\)](#)
10. [^ View all posts in Programming \(www.dancohen.org\)](#)
11. [^ View all posts in Software \(www.dancohen.org\)](#)
12. [^ View all posts in Standards \(www.dancohen.org\)](#)
13. [^ View all posts in Web \(www.dancohen.org\)](#)
14. [^ RSS 2.0 \(www.dancohen.org\)](#)
15. [^ leave a response \(www.dancohen.org\)](#)
16. [^ trackback \(www.dancohen.org\)](#)

Excerpted from *Dan Cohen's Digital Humanities Blog » Blog Archive » Creating a Blog from Scratch,*

Part 5: What is XHTML, and Why Should I Care?

<http://www.dancohen.org/2006/01/05/creating-a-blog-from-scratch-part-5-what-is-xhtml-and-why-should-i-care/>

READABILITY — An Arc90 Laboratory Experiment

<http://lab.arc90.com/experiments/readability>