

1999

# The Offspring of SGML and HTML

Terry Ballard

*Gill Library, College of New Rochelle, tballard@cnr.edu*

Follow this and additional works at: <http://digitalcommons.cnr.edu/gill-publications>

 Part of the [Library and Information Science Commons](#)

---

## Recommended Citation

Ballard, T. (1999). The Offspring of SGML and HTML. *Information Today*, 16(11), 54.

This Article is brought to you for free and open access by the Gill Library at Digital Commons @ CNR. It has been accepted for inclusion in Gill Library Publications by an authorized administrator of Digital Commons @ CNR. For more information, please contact [lfazzino@cnr.edu](mailto:lfazzino@cnr.edu).

## The Systems Librarian

by Terry Ballard

# The Offspring of SGML and HTML

*XML, the easy-to-use, next-generation Web language, shows much promise*

*"Prediction is hard, especially about the future."*

—Yogi Berra

When I worked at another university, I participated in a major project that involved SGML (Standard Generalized Markup Language). SGML was actually the parent of the HTML (HyperText Markup Language) we have come to know and love. SGML had the potential to handle scholarly and detailed information in a way that was much richer and more powerful than HTML. SGML did, however, have several noticeable problems:

- It had countless variations. The one that NASA scientists used was of no use to archivists in law libraries and vice versa.
- It could not be read in a standard browser. Those needing access to databases that used SGML either had to add a special plug-in to their browsers, or they could use programs on the server end that translated data on-the-fly into HTML.
- It was so powerful that it confused even the people who worked with it, including vendors I knew that were in the business of converting data into SGML. OK, it confused me too.
- The people you had to have in your corner were running away from it. At the time I first worked with SGML, Microsoft had begun talking about no longer supporting the protocol—not a good sign.

“  
SGML was  
actually the parent  
of the HTML  
(HyperText  
Markup Language)  
we have come to  
know and love.  
”

Given all of this, I quietly predicted that SGML was not on its way to the Technology Hall of Fame. At the time, however, people were talking about another child of SGML: XML (eXtensible Markup Language), which was to be a compromise format that would boost the ability of the Web to handle complex data. It would also be accessible to people who had cut their teeth on HTML. It sounded promising, but I didn't hear any more about it for a few years.

XML showed up again on my radar early this year at the Computers in Libraries Conference in Arlington, Virginia. It was one of the new developments that people were starting to talk about, so I decided it might be a good time to revisit the subject. First, I went to the Web page of the Graphic Communications Association (<http://www.cga.org>), where it noted that XML has been designed for ease of implementation and for interoperability with both SGML and HTML. The site also explains some of the design goals of XML, which should 1) be straightforwardly usable over the Internet, 2) support a wide variety of applications, 3) be compatible with SGML, 4) enable easy-to-write programs that process XML documents, 5) produce XML documents that are legible and reasonably clear, 6) prepare XML designs quickly, and 7) enable easy-to-create XML documents.

I interpret all of this to mean that XML allows you to add significant power to Web pages, but in such a way that even recycled English majors like me can comprehend it. This contrasts with Java, which I can only work with as long as I don't expect it to make any sense. Interestingly, one of the hot new trends is combining Java with XML.

### Fall Internet World '99

Years ago, when I attended the first Internet World conference, it seemed quite imaginative to hold an entire conference dedicated solely to the Internet. (There were only about five rows of vendors then.) It has certainly changed a lot since then. While looking at the Web site for the recent Fall Internet World '99 conference Web page, I noticed that a number of the vendors listed were working with XML. I also saw that I was eligible for a press pass, so I dropped them a line.

A few days later, the vendors started making contact. Each day, I would get four or five voice-mail messages from Heather, Lisa, Sean, Jennifer, etc., all calling from Park Avenue public relations firms and inviting me to meet the presidents of various Internet start-up companies. That was exciting for about a day, until the phone messages were accompanied by 20 e-mail messages. I soon realized that most of these had nothing to do with library systems, so one message that contained the words "e-commerce solutions" was filtered out. The remaining 10 percent were then carefully examined, and I made a number of appointments with vendors that were involved in some aspect of XML.

With more than 500 vendors, the conference took 2 days just to take it all in. Judging by the number of exhibitors, their willingness to hand out free T-shirts, and the lavish appointments of the press room, I concluded that a lot of money was being made on the Internet. My second observation was that XML was the hottest thing going at the conference. I saw dozens of vendors embracing the protocol—one of them even turned questions about XML

into a quiz show. I was coaxed into being one of the three contestants and came in second on a technicality. The people at Informix Software, Inc. were much more serious about the technology. I spoke with Marielle Valdez-Fong, department manager of the newly formed XML section, and Jeffery J. Menz, executive director of product marketing and management for the Informix Dynamic Server 2000.

“  
Not surprisingly,  
XML will  
be showing up  
in the most  
basic library  
automation tool:  
the online catalog.  
”

Valdez-Fong said that XML is the "lingua franca" of the Internet, meaning, among other things, that it can map relational data to the Web. Menz and Valdez-Fong also told me that both Internet Explorer 5 and Netscape 5 would be able to read the format without using any intervening programs.

Later that day, I saw in a practical sense what this means. The people at the Borland booth were showing off their new relational database product, Delphi 5. They brought up a Web display with a standard four-column table with names, addresses, regions, and sales figures. Using XML, the user could sort the data as if it were in the relational database. This means taking a Web page far beyond a flat display of data to interact in a useful way.

Another sign of XML's future is the fact that Microsoft has an XML-development division. You can find out more at <http://msdn.microsoft.com/xml/default.asp>. Not surprisingly, XML will be showing up in the most basic library automation tool: the online catalog. I spoke with Steve Silberstein, vice president of Innovative Interfaces, and found out that his company plans to use the language in future releases of its Web-based OPACs, although it is not slated for their 2000 release. It might be particularly useful in the company's new generation of Windows-based acquisitions modules.

In addition to the previously mentioned <http://www.cga.org> page, I found that <http://www.xml.com> is also a good jumping-off point for things such as XML-compliant browsers and authoring tools.

### Mixing Technology with Pleasure

After all of the enlightenment at the recent Internet World, naturally we were hungry and thirsty. So the invitation we got from One Voice Technologies was particularly appealing: an evening cruise of New York Harbor with dinner and cocktails. As the boat slid under the Brooklyn Bridge, the guests were invited to see the company's demonstration, if they wished. One Voice Technologies produces a voice-recognition system that allows the user to give orders to the computer, such as "Show me the names of New York hotels with rooms available next week." An animated figure will say "Yes, master" and bring up the results. My wife and I immediately saw the possibilities of this for visually handicapped library patrons. Afterwards, the boat parked at the foot of the Statue of Liberty, which was a beautiful sight on a perfect moonlit night. It made me think of a time 10 years ago, when I was still finishing up my library degree while working as a library paraprofessional in a dead-end public library position. These new technologies have been very good to me.

Terry Ballard is the automation librarian at Quinnipiac College in Hamden, Connecticut. He can be reached at [ballard@quinnipiac.edu](mailto:ballard@quinnipiac.edu), or through his Web page at <http://www.geocities.com/Athens/Delphi/363>.

### EOS International

(continued from page 53)

are also available from EOS International for more complex reporting requirements.

According to the company, the Q Series is a full-featured library information management system providing sophisticated information retrieval with advanced searching options that overcome the limitations of Boolean searching. Q Series does this by providing additional tools, such as natural-language querying, "hypersearching," concept searching, and pattern matching. Through these tools and other features, the Q Series provides a single point of access to information. This information can include a library's own collection, a branch library's collection, multimedia files, CD-ROMs, URL links, and commercially or internally published documents found on the Internet, intranets, and extranets. According to the announcement, Q Series was nominated for the *Computerworld* Smithsonian Award for its innovative use of information technology and its ability to make information more manageable and accessible to users.

The Q Series is a client/server-based system. Client workstations require a 100-percent compatible PC processor running Microsoft Windows 95, 98, or NT Workstation 4.0. The Q Series 2.0 server, using SQL Server, runs on Windows NT version 4.0 or Sun Solaris version 2.6 running on a SPARC or Sun Enterprise server.

Source: EOS International, Carlsbad, CA, 760/431-8400; Fax: 760/431-8448; <http://www.eosintl.com>.