
Subject: Re: Another XML Question

Posted by [Michael Wallace](#) on Thu, 24 Mar 2005 00:37:17 GMT

[View Forum Message](#) <> [Reply to Message](#)

> Oh, of course. Now what the hell is a DTD again?

Document Type Definition.

It all begins with SGML, Standard Generalized Markup Language, which is what other markup languages (XML, HTML, XHTML, etc) are based on. SGML defines the basics of markup and sets a few rules, but not many. XML was created by the W3C to be a simplified version of SGML suitable for sharing data over the Internet. There is nothing that limits XML to Internet-only work -- it just happened this was the original audience, but it has now grown into a number of different worlds. Anyway, XML simplified SGML by adding some extra structure, but didn't limit the possible tags someone could define.

Even though XML is simpler, it could still be really ugly to work with because the tags used are arbitrary. It'd be difficult to try to work with a file without some prior knowledge about acceptable tags and attributes, as our friend David has discovered. To get around the problems of arbitrary definitions, a DTD is used. A DTD simply defines all of the acceptable tags, nesting rules, acceptable attributes for given tags and so on. Having such a definition allows data to be interpreted and transferred much easier as you don't have to guess at what's a valid tag or attribute and don't have to guess about what comes next. There are plenty of XML parsers which only need a valid DTD and they can read any XML file corresponding to that DTD. It's much easier than rolling your own each time.

And if you ever here Schema or XSD mentioned along with XML, that's a reference to the successor of DTD, but DTDs are still in high use, so it's worthwhile to know how they work. Plus, it's easier to learn DTDs first and then pick up XSD. Anyway, I saw another message in this thread with an example of a DTD, so I won't bother with throwing up another one. Also, one of the best resources I've found for getting started on stuff like this is the W3C's www.w3schools.com site. There's a lot of good info on XML in all it's variations.

-Mike
