

Freely available structures: XML Document Type Definitions you can use today

This article discusses a few of the many freely available XML Document Type Definitions available to assist authors in creating valid XML content

Kay Ethier

Scott Abel



rganizations of all sizes are beginning to realize how content and its reuse across the enterprise can improve productivity – and the bottom line. The need for change is driven by the desire to better manage information assets (documents, creative ideas, illustrations, charts, graphics, multimedia, etc.) and eliminate costly processes that fail to facilitate the effective and consistent re-use of content. At the heart of managing content for re-use however lies the job of exposing the underlying structure of that information.

Organizations of all sizes are beginning to realize how content and its reuse across the enterprise can improve productivity – and the bottom line

The most significant way that structured documents differ from unstructured ones is that structured documents include “rules.” These rules formalize the order in which text, graphics, and tables may be entered into a document by an author. For example, in an unstructured document, a paragraph has specific formatting – font, size, and spacing. In a structured document, this same paragraph also has an exterior “wrapper” that governs the elements that are allowed to appear before and after it. The rules of these elements are often defined in a Document Type Definition (DTD).

XML is meta markup language that is used to create new markup languages. It’s most commonly used (in the pub-

lishing arena) to create tag sets and processing instructions (DTDs) that describe structured content for presentation in text documents, but it can also be used to describe, manage, and deliver content of all types (text, images, voice, forms, multimedia files, and so on) and to transform transactional data between disparate database systems.

To learn more about XML and whether it’s the right choice for your organization, read, XML: The Answer To Everything? (http://www.freesoftwaremagazine.com/free_issues/issue_01/focus_format_xml/), Free Software Magazine, Issue 1.

Darwin Information Typing Architecture (DITA) DTD

Managed by OASIS (Organization for the Advancement of Structured Information Standards), the Darwin Information Typing Architecture (DITA) 1.0 standard was released for public review in early 2005. DITA is designed for content creators, providing a framework for easily authoring online help, user guides, website content, and more. Content creators who use DITA author content in a topic-based tree, with information “typed” by using different subelements. The four main elements are: topic, task, concept, and reference. The real beauty of DITA is that it provides an architecture - unlike other DTDs - that allows for easy expansion through specializations (customizations to the DTD). A schema is also available for DITA.

The Adobe FrameMaker DITA Development Group



The DITA DTD, schema, and related files including XSLT, are available in a ZIP file (<http://www.oasis-open.org/committees/download.php/11428/cd1.zip>).

There is also an EDD (<http://groups.yahoo.com/group/framemaker-dita>) (element definition document) and related files, which works with Adobe FrameMaker.

Learn more about DITA here (<http://www-106.ibm.com/developerworks/xml/library/x-dital/index.html>).

DocBook XML DTD

Available by download from many places, and installed with many tools as a “freebie”, the DocBook XML DTD gives users absolutely every element they might need to create software or hardware documentation. However, DocBook may not be the right choice for some authors, especially those unfamiliar with authoring structured XML content. Additionally, authors have complained that the number of elements contained in DocBook is overwhelming (over 300 valid elements are included) and that as a result, using the DTD is more difficult than it ought to be. For that reason, some organizations have adopted the Simplified DocBook DTD (detailed below).

Check out SourceForge.net and other sites for DocBook projects, and downloadable DTDs and sample files. You may also try a Google search for “Norm Walsh and DocBook” to find examples and historical data on this

DTD...it has pre-XML roots in SGML (Standard Generalized Markup Language).

Learn more about DocBook here (http://sourceforge.net/docman/?group_id=21935).

The real beauty of DITA is that it provides an architecture - unlike other DTDs - that allows for easy expansion through specializations (customizations to the DTD). A schema is also available for DITA

A downloadable structured Adobe FrameMaker DocBook XML example project created by Steve Whitlatch is available online as well. It includes an expanded discussion of using the structured version of Adobe FrameMaker and the DocBook XML DTD and details some new auto-formatting features for tables, graphics, and xrefs. Additional information about the DocBook Extensible Stylesheet Language (XSL) project is also provided.

Learn more about DocBook for Adobe FrameMaker here (http://www.getnet.net/~swhitlat/DocBook/docbook_section.html).

Simplified DocBook XML DTD

According to DocBook.org, the Simplified DocBook DTD “is a small subset of the DocBook XML DTD” (105 elements) created to help authors who may be overwhelmed by the sheer number of valid elements in the original DocBook XML DTD (over 300 elements).

Learn more about the Simplified DocBook XML DTD here (<http://www.nwalsh.com/docbook/simple/index.html>).

A variety of DocBook resources are available here (<http://www.dpawson.co.uk/docbook/reference.html>).

DocBook Slides DTD

The DocBook Slides DTD is also a free DTD contributed for public use by Norm Walsh. By using the DocBook Slides DTD (and the many stylesheets and support files

The DocBook project



that come with it), you can immediately begin creating XML slides and transforming them into online presentations. The main element is the “Foil” – referring to the old, plastic floppy slides that used to be used for presentations. Users may create a “FoilGroup” and place many “Foils” within their presentation. At the start of a presentation there are elements for creating title and speaker information as well. The DocBook Slides stylesheets transform each foil into an online page with the user’s choice of look-and-feel, a table of contents, and complete navigation (next/previous/contents).

Visit SourceForge.net to download the DocBook Slides project files and corresponding DTD.

Learn more about DocBook Slides here (<http://sourceforge.net/projects/docbook/>).

A DocBook Slides demonstration is available here (http://sourceforge.net/project/showfiles.php?group_id=21935&package_id=59365).

U.S. Military DTDs

Recently, the Department of the Navy (DON) released its Naming & Design Rules (NDR) version 2.0 for creating and working with XML structures. Version 2.0 firmly establishes a common structure and language for authoring and presenting content, and enables the discovery and reuse of common content elements across the naval enterprise.

Learn more about the U.S. Navy’s XML NDR here ([https://www.doncio.navy.mil/\(b5stca55srias320waniyi45\)/storefront/Product_Detail.aspx?ID=1318](https://www.doncio.navy.mil/(b5stca55srias320waniyi45)/storefront/Product_Detail.aspx?ID=1318)).

Learn more about the U.S. Navy’s information management plan here ([https://www.doncio.navy.mil/\(b5stca55srias320waniyi45\)/fy04stratplan/](https://www.doncio.navy.mil/(b5stca55srias320waniyi45)/fy04stratplan/)).

The U.S. Department of Defense (DOD) maintains an XML Gallery – a metadata registry and clearinghouse for military XML information. The purpose of the Gallery’s XML Registry is to promote the consistent use of XML, both vertically within individual agencies and horizontally across all DOD agencies.

By using the DocBook Slides DTD (and the many stylesheets and support files that come with it), you can immediately begin creating XML slides and transforming them into online presentations

Learn more about the U.S. DOD XML Gallery (<http://xml.dod.mil/xmlreg/user/information.cfm>).

Conclusion

XML DTDs allow you to add levels of control to the documents you create. DTDs facilitate structured content development, content reuse, consistency, and interoperability. Whether you create your own XML DTDs or utilize one or more of the many freely available XML DTDs is up to you. No matter which approach you take, make sure you learn everything you can about your business needs (those of the end users, as well as the needs of your content creators) before you decide which route to take. Additional resources are provided below for those who’d like to learn more about XXML DTDs.

Bibliography

Document Type Definition Tutorial (<http://www.w3schools.com/dtd/default.asp>) – a free online class from w3schools

Fundamental Concepts of Reuse (<http://www.managingenterprisecontent.com/>) – under-

standing content reuse (Ann Rockley, Pamela Kostur and Steve Manning)

XML Schema Tutorial (<http://www.w3schools.com/schema/default.asp>) – a free online class from w3schools

XML Weekend Crash Course (<http://www.wiley.com/WileyCDA/WileyTitle/productCd-0764547593.html>)

Copyright information

© 2005 by Kay Ethier, Scott Abel

(The following license is effective immediately)

This article is made available under the “Attribution-NonCommercial-NoDerivs” Creative Commons License 2.0 available from <http://creativecommons.org/licenses/by-nc-nd/2.0/>.

About the author

Kay Ethier is an Adobe Certified Expert in FrameMaker 7.x and several prior versions. She instructs training classes, performs consulting, and provides support to clients in a variety of industries. Kay resides in the Research Triangle Park area of North Carolina and works for Bright Path Solutions (<http://www.travelthepath.com>). In 2001, Kay co-authored the book *XML Weekend Crash Course* (Wiley/HungryMinds). She has most recently been a contributing author on *Advanced FrameMaker* (TIPS Technical Publishing) and *XML and FrameMaker* (Apress).

Scott Abel is a technical writing specialist and content management strategist whose strengths lie in helping organizations improve the way they author, maintain, publish and archive their information assets. Scott has considerable experience in XML-based enterprise content management strategy, information architecture, software tool selection, high tech marketing communications, business process analysis, technical writing and editing, usability, content localization, and designing single source solutions.