

Publishing PDFs from DITA

Introduction

- Scott Prentice, President of Leximation, Inc.
- Specializing in FrameMaker plugin development as well as structured FrameMaker conversions, consulting, and development. FrameMaker user/developer since 1991.
- Developed DITA-FMx, a FrameMaker plugin for efficient DITA authoring and publishing.
- Developer of custom Help systems and creative/functional web applications.
- Coined the term "AIR Help" in 2007 after learning about Adobe's new AIR technology.
- Interested in creating innovative ways to provide user assistance that is actually used.

What is DITA (briefly)

- DITA - Darwin Information Typing Architecture
- An XML format for authoring in a topic oriented structure
- Highly reusable and modular
- DITA topics can be easily rearranged and reused for different deliverable types
- Filtering/conditionalizing based on attributes
- Tagging is applied in a semantic manner rather than to achieve specific formatting

Publishing from DITA

- Content authored in DITA can be easily published to various output formats with the DITA Open Toolkit
- The DITA OT makes it very easy to experiment with various output types, but it is really just the starting point
- Other tools are available for publishing from DITA, which can make the process easier than the OT, but you end up relying on proprietary tools
- Producing PDFs from DITA requires a bit of analysis and some hard work

Popular DITA to PDF Publishing Tools

XSL-FO processors:

- Apache FOP
- Renderex XEP
- Antenna House Formatter
- Quadralay ePublisher

Proprietary rendering engines:

- SDL XPP
- DITA Exchange OpenXML Publishing
- MadCap Blaze/Flare
- Adobe FrameMaker / Leximation DITA-FMx

What about XSL-FO?

- XSL-FO (XSL Formatting Objects) is a markup language for formatting XML documents
- Typically used for generating PDFs
- Part of the W3C's XSL specification (also XSLT and XPath)
- Typical workflow for converting XML to PDF using FO:
 - User writes content in XML
 - XSLT transformation generates the XSL-FO
 - XSL-FO is passed to the FO processor to generate PDF

Benefits and Limitations of XSL-FO

Benefits

- Low entry cost
- XSL-FO is an XML language
- FO markup language is largely based on CSS
- Designed to work for all written human languages

Limitations

- XSL-FO coding is expensive and complicated
- Tables that span pages may not break as expected
- No way to render elements on one page in relation to another page
- Difficult to apply balanced vertical spacing on a page
- Processors may use extensions to implement features, so the FO stylesheets may not be portable between processors

Apache FOP

- Open Source / Free
- Java application, runs on "any" platform
- Default PDF output option provided with the DITA-OT
- Rendering engine uses XSL-FO
- Requires XSL-FO developer to maintain stylesheets
- Supports popular raster and vector graphic formats, + SVG
- Can be automated/scripted
- <http://xmlgraphics.apache.org/fop/>

Renderex XEP

- \$360 standalone, \$4400 server
- Java application, runs on "any" platform
- DITA processing through the DITA-OT
- Rendering engine uses XSL-FO
- Requires XSL-FO developer to maintain stylesheets
- Renderex Visual XSL - GUI for generating XSL-FO stylesheets for forms-based output
- Supports popular raster and vector graphic formats, + SVG
- Can be automated/scripted
- <http://www.renderx.com/tools/xep.html>

Antenna House Formatter

- \$1250 standalone, \$5000 server
- "Lite" (under 300 pages) \$400 standalone, \$2000 server
- Java application, runs on "any" platform
- DITA processing through the DITA-OT
- Rendering engine uses XSL-FO and CSS
- Requires XSL-FO developer to maintain stylesheets
- Supports popular raster and vector graphic formats, + SVG
- Supports MathML
- GUI supports over 50 languages
- Can be automated/scripted
- <http://www.antennahouse.com/product.htm>

Quadralay ePublisher

- \$1999 perpetual, \$795 one year subscription
- Windows only
- Processes DITA, FrameMaker, and Word documents
- PDF rendering engine uses XSL-FO
- Uses Apache FOP as the default formatting engine, but other processors can be used
- Provides GUI for creating FO stylesheets
- Converts DITA, FM, and DOC to WIF, then to PDF via FO
- Exports other online formats: HTML, Wiki, Help, ePub
- Automated publishing requires Automap
- <http://www.webworks.com>

Other FO Processors

Lunasil Xinc

- Java-based processor
- \$95 developer, \$495 server
- <http://www.lunasil.com/> (website last updated in 2005?)

Altsoft Xml2PDF

- Windows application (.NET-based application)
- \$55 workstation, \$1495 server, plus other options
- <http://www.alt-soft.com/>

PTC Arbortext Publishing Engine + Arbortext Styler

- Java-based server application
- Enterprise publishing solution
- <http://www.ptc.com/products/arbortext/publishing-engine/>

SDL XPP

- \$50K to over \$200K depending on options
- Windows, Solaris, and Linux platforms
- Rendering engine uses proprietary formatting language to support high quality layout and typography needs
- Provides hyphenation and justification in over 30 languages including Chinese, Korean and Japanese
- XyView client creates layout and formatting templates
- Full editing of composed output through XyView client
- Typically automated but can be run interactively
- <http://www.sdl.com/en/xml/products/sdl-xpp/>

DITA Exchange OpenXML Publishing

- Pricing not finalized
- Windows only
- Uses proprietary process to convert DITA to OpenXML
- OpenXML is opened in Word (2007 or 2010) and specified template is applied for formatting
- Client application available for interactive conversions
- Server-based processing via SharePoint enables workflow, document reviews, digital signatures, etc.
- Input can be a mix of DITA and DOC files
- Also exports to PowerPoint, one slide per topic
- <http://dita-exchange.com/>

MadCap Blaze/Flare

- \$699 Blaze, \$999 Flare
- Windows only
- Imports DITA, FrameMaker, Word, and other formats
- Exports to PDF and numerous other online formats
- Uses proprietary PDF rendering engine
- Provides GUI for defining layout/formatting
- Automated publishing via a command line interface
- <http://www.madcapsoftware.com/products/flare/>

Adobe FrameMaker

- \$999 workstation, \$7999 server
- Windows only
- Proprietary layout/formatting through template and EDD
- Native PDF export or through Adobe Distiller
- Automated publishing using scripting or FDK development
- Standard FM “build” process:
 - Develop templates and EDDs
 - Save DITA map to FM book and chapter files
 - Set up generated lists (toc, index, etc.)
 - Apply formatting and pagination properties
 - Save book to PDF
- <http://www.adobe.com/products/frameMaker/>

Adobe FrameMaker + Leximation DITA-FMx

- FrameMaker (versions 7.2, 8, or 9) + \$185
- Windows only
- Proprietary layout/formatting through template and EDD
- Native PDF export or through Adobe Distiller
- Automated publishing with Auto DITA-FMx
- Standard DITA-FMx “build” process:
 - Develop templates and EDDs
 - Configure for automated creation of generated lists (toc, index, etc.) and component formatting / properties
 - Save DITA map to FM book and chapter files, then save to PDF
- <http://leximation.com/dita-fmx/>

FrameMaker is Ideal for High Quality PDF Output

- Relatively easy to define the layout and formatting
- Easy to make minor adjustments when needed
- Ability to modify the pagination/formatting after composition but before PDF creation
- Likely to have in-house expertise
- Easily set up templates for multiple languages
- Likely to be less expensive in the long run
- Can use FrameMaker for PDF generation even if you're using another tool for DITA authoring
- Save time and ensure consistent output with DITA-FMx

Choosing the Right Tool

Determine your PDF needs

- Online and/or print PDF deliverable?
- One layout or many?
- Frequent changes to formatting/layout?
- Possibly want to make minor adjustments to the output?
- High quality typographic requirements?
- Large output volume?

Assess your available resources

- Do you have an XSL-FO developer on staff?
- Do you have in-house expertise with FM template/EDD development?
- Available contract developers?

Wrap Up

Resources:

- http://gilbane.com/gilbane_report.pl/94/XSLFO_Ready_for_Prime_Time.html
- http://en.wikipedia.org/wiki/XSL_Formatting_Objects
- <http://www.sagehill.net/docbookxsl/FOprocessors.html>

Questions?

- Scott Prentice <info AT leximation DOT com>
- **Leximation** - <http://www.leximation.com>