Publishing Equations? Do the Math(ML)

By Robert J. Boeri - Aug/Sept Issue, Posted Sep 16, 2003
http://www.econtentmag.com/?ArticleID=5035

If you work with STM (scientific, technical, medical) publishing, sooner or later you'll need to produce mathematical expressions beyond the reach of a QWERTY keyboard. Until recently, the solution was simple: Use the equation editor in your standard word processor or a technical publishing system. Simple, that is, until you try to bridge the gap between authors and production, especially if you need to deliver multiple versions—Web, PDF, print, etc. Nearly 10 years ago, I wrestled with transferring technical content between authors and production and built a solution using WordPerfect styles and a Xyvision—now XyEnterprise—publishing system. It worked, but equations were always a problem. There was no way to transfer equations between WordPerfect and Xyvision other than converting them to graphics. Keeping track of hundreds of equation graphics is difficult at best.

As XML has transformed the content landscape, it's now changing the way to express math equations. The World Wide Web Consortium (W3C) developed two standards garnering vendor support: MathML (Mathematics Markup Language) and SVG (Scalable Vector Graphics). SVG, originally standardized in 2001, is XML-based, allowing SVG graphics to be manipulated like other XML content. MathML, initially released in mid-1999, aims to solve several problems in expressing and using equations: Notation (equation symbols), content (the meaning of the equations) to facilitate machine processing, and ultimately the searching and indexing of these symbols. These lofty goals don't even address the problem of displaying equations in browsers or print.

My rule of thumb is that it usually takes vendors two or three years to develop products implementing promising standards, and that rule applies today. Adobe Illustrator is SVG-capable, and Corel has recently released its Smart Graphics Studio for building SVG applications. On the MathML side, vendors like Design Science deliver useful products that make a compelling case to begin planning integration of MathML in your publishing processes.

Benefits of migrating to MathML are obvious. Using MathML to express equations in an XML publishing system means that your entire documents can be XML, rather than XML with links to equation graphics. MathML lets your systems generate graphical views of equations at the end of your publishing process, where you also produce Web, PDF, or other outputs. Implementing MathML will remove the need for keeping and managing large numbers of equation graphics throughout the publishing lifecycle and frees you from managing inevitable changes to those graphics throughout the process. By using MathML, you preserve your content investment as growing numbers of vendors provide ways to derive added value from MathML investments. Such added value could include searching equations (you can't do that with graphics) and allowing users to copy and paste equations directly into math processing systems such as those from Wolfram Research's Mathematica.

I asked several vendors about the benefits of MathML, including Paul Topping from Design Science. Topping worked with the W3C and contributed his expertise to help develop MathML. Design Science develops and markets MathType, an interactive tool for Windows and Macintosh for creating mathematical notation for word processing, Web pages, desktop publishing, and for STM documents. MathType is the professional version of the Equation Editor in Microsoft Office and many other products. Design Science also offers a free browser plug-in for viewing MathML equations. Topping said that STM publishers today are interested "in using MathML as part of an overall XML workflow. They've bought into the idea of keeping as much of their content as possible in XML and, as MathML is the XML language of choice for math, MathML is a natural part of that solution." Don't go looking for MathML support in word processors or most STM publishing systems just yet, although Arbortext has begun implementing MathML in a partnership with Design Science.

Other vendors like SchemaSoft offer tools to convert MathML to SVG, for displaying equations in browsers with an

SVG plug-in. According to Corel, which like Adobe has developed rich support for SVG, 70% of browsers have an SVG plug-in. Without control over the plug-ins your customers use, SVG renditions of equations make sense. Lastly, consider Mathematica. Wolfram's spokesman Pavi Sandhu told me that users can copy MathML equations from a Web browser for use in Mathematica, letting Mathematica users evaluate the equations. This capability can provide a value-added service to your publication subscribers who also use Mathematica.

Vendors are exploiting SVG in far broader ways than just displaying math formulas, including an alternative to Acrobat PDF files. Read my next column for more information about Acrobat 6 and SVG alternatives.