February 2016

Simply XML February 2016 Newsletter

This is the Simply XML Newsletter. We welcome the new folks from the recent TC Camp in Santa Clara. We hope you enjoy these newsletters, but if you want to opt out, please unsubscribe at the bottom of the page. We would love to hear your comments and you can email us at hello@simplyxml.com.

Have a great day!

Myth Busters

by Doug Gorman

Three Historical Assertions

1. Man will never walk on the moon.
2. DITA is too complicated and will never be easy to implement.
3. You can't get well-formed, valid DITA from Microsoft Word.

Let's take these one at a time.

1.   Man will never walk on the moon.

I saw this on television at a time when you couldn't fake this stuff with digital animation. In 1969 Neil Armstrong took "one small step for a man, one giant leap for mankind." He walked on the moon.

Result: Myth busted on July 20, 1969.

2.   DITA will never be easy to implement.

Well first of all, DITA is complicated, but so are machine language, cell phones, and television. Human problems are complicated, but DITA is not really a human problem. DITA is a mark-up language and like machine language, cell phone technology, the technology of television, DITA usefulness comes from a different level. It comes from the ability to hide the technology with an easy to use layer. It's all about what you can do with XML when it is underneath.

Out of the box, including specializations there are some 500 or more elements and a host of structures and facilities that someone who wants to "know DITA" needs to know (or thinks they need to know).
Now enter Lightweight DITA which offers an easy to learn incredibly useful facility for the rest of us. It constrains DITA to DITA Topic and DITA Map.

Most people have an innate understanding of what a Topic is—a unit of content that can stand on its own. And like most structured content, DITA Topics contain paragraphs, tables, lists, and images – the structures that most educated knowledge workers create every day in the published structures that information consumers require.

A DITA Map is merely a collection of DITA Topics that contribute to a larger information need or set. It is not so hard to think about this. Most of us have files (sort of maps) by category that contain related articles or writings on that subject. You've been using file structures for years.
But now with a Lightweight DITA architecture, a shared repository, and modern metadata, and a broad range of publishing capabilities are readily accessible.

You can use that smart well-prepared guy's Topic in your DITA Map.

- With permission, you can search through other people's Topics to find content you need to use.
- You can delight your audience with PDF, HTML, EPub, or mobile content.
- And you can easily contribute your brilliance specifically to the organizational effort.

Better still, you can rely on your knowledge and experience without the confounding requirement of learning either DITA or XML.
Result: Myth busted by newly refined approaches to DITA.

3.   You can't get well-formed, valid DITA from Microsoft Word.

Last weekend at Liz Fraley & Company's TC Camp in Santa Clara I spoke with at least 5 people who were absolutely, 100%, totally and unequivocally certain that they had to dump Microsoft Word in order to implement DITA or another XML structure. Reinforcing this understanding was the knowledge that they could not allow their organization to operate in an environment of mixed tools. Clues contributing to this knowledge include the following:

- You can do pretty much anything in MS Word. There are no constraints.
- Word-based writers will always spend around half of their time making their content look "good."
- Review functions and track changes are the gold standard for content review. AND
- There will always be employment for intermediate editors to turn Word-based content into valid XML.

Content Mapper is an authoring tool that takes full control of the Word User Interface. It uses style sheet transforms so that the author sees Word, but the repository sees XML. Content Mapper presents Word, but in a constrained way that is driven by the document rules, technically called a schema.

- The author can create a list but doesn't get to choose wingding squares as his bullets.
- The author can reuse other people's content, but that content already exists under an information standard
- The author can invoke publishing look and feel locally or from the server and the result will be a look and feel that is approved and looks "good" at a universal level.
- The author can try to trick the system, but cannot save invalid XML.

Result: Myth busted because with a product like Content Mapper, you can easily get well-formed valid XML from Microsoft Word.

Conclusion:

Microsoft Word is an authoring tool used by a billion people. As an add-in to

Microsoft Word, Content Mapper with Lightweight DITA will allow organizations to

make XML and DITA a realistic enterprise architecture for content. Organizations can

implement this important architecture to get control of Microsoft Word and to

promote reuse and flexible publishing to the delight of the people who really matter –

your information consumers.

Content Review-- It's Complicated

By Doug Gorman

Content Review is Complicated

It's really great to go to conferences because you get to have really interesting, often intelligent discussions with great people about interesting subjects and problems. Such has been the case over the last year when talking about the difficult issue of

content review.

The only conclusion I've been able to reach from the vendors is, "Where you sit is where you stand." The only conclusion I've been able to reach from organizations is, "No one has actually figured it all out yet." The reason is that there are lots of problems and just as many solution approaches-- it is complicated!.

The Standardization Problem

Back in the 1800's Eli Whitney invented the concept of standardized parts for manufacturing. Back then the focus was on fighting battles with guns, but today's focus is on fighting battles with words, tables, charts, and images. Yes the craftsmen who produced guns in the 1700's were replaced by the craftsmen who produced dies and other methods for making guns quickly and efficiently. The same is true for our knowledge workers today. Individual writers are now being asked to contribute to a more efficient and effective systems we call the content supply chain. Here's the problem: how can quality happen when our content does not fit together? And how can reviews ever move the quality and consistency of information forward when structures are ill-defined and incomplete i.e. not standardized?

The Technical Review Problem

There is no common technical standard for the review functions themselves. In DITA, all the vendors seem to hide their review functions in Processing Instructions but those functions are represented diversely. It is like saying that "we are all the same and compatible if we live in houses on earth. All books fit together because they are made up of words."

The Technology-based Review

Technology-based tools can help. MS Word can help with spelling and some grammar issues. Readability scales can help you understand the educational level of the document. Acrolinx can take you to another level of content standardization and help improve content across cultures and languages. On the Flesch-Kincaid reading Ease scale this article gets a score of 59.5 and the average grade level of 5 applied tests or reading ease is 10.2. (I hope many of you completed high school. J the previous article was a 64.3 and grade level of 8.7, perhaps because I had a clearer idea of what I was talking about.)

Document Size and Timing Problem

Some want the whole document to review thereby thinking about context and document consistency. Some are hurried and harried and just want the Topics that require their review. Some are forced to take a whole document, say 100 pages and spend the majority of their time on the first 10 pages, skimming the rest as the deadline approaches. And then there's the "You've had it for two weeks and I need it tomorrow variation of this." And let's not forget the "my boss measures me on this, but my reviewers are not measured or motivated to do a good job."

The Type of Review Problem

Often dependent on purpose which varies. this problem can result from the variety of review objectives, for example:

- Give me a quick assessment of whether I'm on the right track.
- The government requires this type of detailed compliance review.
- Legal requires that every word must conform to government and legal standards.
- Grammar should reflect proper English (French, Mandarin, etc.)
- Finally, there's "the Boss is worried about the political implications review" due to new ground covered or organizational implication.

The Concurrency Problem

The whole thing can be complicated by the concurrency problem. In the good ole days everybody got a copy of the Word document, used track changes and comments and emailed the document to the author who created another draft after considering the various comments and suggestions. With CMS's, check-in and check-out it's unrealistic to check out a large document and take your time reviewing. It is similarly unrealistic and suboptimal to let everyone do their reviews independent of other people's suggestions.

From Simply XML's Perspective

Content Mapper supports MS Word's review function and I will state with certainty that the reviewer actions can be the most varied in Word. You can add, delete, change, move, restructure to your heart's content. Similar to what we do with Word-based XML authoring, we confine review comments to the valid UI functions that are supported by the schema.

When we are working in a mixed environment, for example with Oxygen as another editor, we can round trip review comments from Content Mapper to Oxygen, but we must leave out the review functions/comments that Oxygen does not support, putting them back in upon return to the repository.

Browser-based Review

This sounds great and everyone with a connection to the web can access the content, but changes are limited. And what happens if I want to work offline.

XML Editor-based Review

Each editor has a different way of supporting review. The UI's are different and they don't work very well with Word-based or Browser-based reviews. But the use of these tools is generally limited to technical people. Oxygen has done a good job of

integrating browser-based and client-based authoring and review, but what about the bulk of authors who use MS Word?

PDF Review

When you use a PDF review tool, you have a common agreed process, but may sacrifice from a less complete review of content, structure, organization, and purpose. And rationalizing those comments in the source document is difficult, at best.

Other Thoughts

It seems that the most appropriate level for group review in a modern organization is the Topic. Don't check the content out until you are ready to do the review. Check it back in with your comments quickly. Go back to the Topic later if you want to critique other comments, or stand on the shoulder of giants who have thought of things that you did not. At the end of the day, once again, information consumers will be delighted.

Conclusion

Like I said earlier, it's complicated. Let's work together this year to figure this one out.

Simply Tips and Tricks

All of our newsletters provide quick tips and tricks to help you simplify the XML world and beyond. Our tip for February focuses on what we know that we know.

In Life:

With life and changing times it is often true that what we knew was known and true because of the context in which we validated this assertion.

In XML:

As the pace of change increases, take some time to validate current assertions based on current states of knowledge and understanding. The light at the end of the tunnel might just be the solution that you are looking for.

There's a reason we call it Simply XML.

Simply XML provides simple, easy to use solutions for creating and publishing XML content. Creating, managing, and publishing content with our XML editor, Content Mapper, minimizes costs and ensures efficiency across your organization. Learn more at www.simplyxml.com.

> ### Request a Free Trial of Content Mapper

Simply XML, LLC.
Tel: +1 (781) 801-9255
Email: hello@simplyxml.com
www.simplyxml.com