

Casting a critical eye on the Next Big Thing in technical publishing.

Web 2.0: The Tipping Point for XML

BY SARAH O'KEEFE, Associate Fellow

By now, you've seen the articles about XML, separation of content and presentation, programmatic enforcement of content organization, structured authoring, and all the rest. You've considered XML publishing, but implementation looks difficult and expensive, and your current workflow is in reasonable shape. So you've been waiting for a compelling reason to make the transition.

That compelling reason has arrived in the form of online user-generated content, otherwise known as Web 2.0. User-generated content isn't new—if we include scribbling notes in the margin of a book. But today, instead of creating annotations for personal use only, we can share information on the Internet with friends, family, or the world using blogs, wikis, forums, and other Web 2.0 applications.

User-generated content is going to be hugely disruptive for technical communication. Your content strategy needs to include Web 2.0, and XML provides a platform to support the seamless integration of "professional" content with "user" content.

Why XML with Web 2.0?

Most of our clients have cited utilitarian reasons for implementing XML—they

expect cost savings in localization or from automated document production.



Increased efficiency and reduced costs tend to make for compelling return on investment (ROI) justifications. Lately, though, we're seeing a new trend—clients want

XML because of the possibilities that structured content opens up. In particular, the integration of corporate content—information created by technical communicators—and user-generated content is becoming a priority.

In an XML-based workflow, content is stored without formatting. When information is published, formatting instructions are added on to the content to produce the final deliverable in print or online. As a result, global formatting changes are relatively easy. The metadata (usually attributes) stored inside the XML content allows for searching and filtering of the information. And finally, the text-based storage format makes it easy to process information. Collectively, these features make it possible to deliver customized information to individual readers. And furthermore, you can mix the official, corporate information with content contributed by end users on forums, wikis, and the like.

Enough with the Paradigm Shifts!

I've written and spoken (a lot) about how the shift from desktop publishing to XML and structured authoring represents a paradigm shift. I expect structured authoring to be around for quite a while, but I'm already seeing signs of a paradigm shift in another area. And I suspect that the new paradigm shift will actually be even more disruptive than structured authoring has been.

Our publishing model—for books, radio, television, and other media—has been a one-to-many model. Only a privileged few can create and distribute content. But today, the cost of distributing content has dropped to zero for anyone with Internet access. In addition, a passionate user's site can easily achieve higher visibility than the official documentation. We now live in a world of many-to-many publishing—anyone can create and distribute information; the content being created by corporate technical writers is just one source of information.

What Does This Mean for Technical Communicators?

There is often a gap between professionally created technical publications and end user needs. The reasons for this include the following:

- It takes a long time to update the official documentation because of lengthy review, approval, and publishing processes. In some regulated industries (such as aerospace), these processes are non-negotiable. However, the end users need the updates, and they need them *now*, so they

find ways around the official processes. Sometimes this means consulting the local expert to find out how things really work. Or perhaps a resourceful user might print out a copy of the official documentation, edit the paper copy, and then put the marked-up version on the bulletin board in the service area.

- In the rare cases where annotation capabilities are provided to end users, they are limited. As a result, users create annotations elsewhere, such as a user forum or mailing list.
- The documentation is not sufficiently candid. It lists only the features and forces the readers to infer limitations. From an end user's point of view, describing limitations of a product, as you often see in reviews, can be very useful. Because documentation tends to avoid such negative assessments, readers go elsewhere for analysis of what a product really can and cannot do. If the gap between what end users need and what the product owner provides is sufficiently large, the end users will find other sources of information. To retain readers, and give the product owner at least some control over the discussion, technical communication departments need to take the following steps:
 - Accelerate document publishing so that information is available when the product is released, and updates are incorporated quickly and accurately.
 - Provide content online in a format that allows for comments and discussion.

- Be explicit about product limitations and workarounds to build credibility.
- Provide valuable content. For example, consider a password-setting feature. The step-by-step instructions on how to set a password are needed and constitute the minimal baseline for documentation. Beyond that, any length requirements (minimum or maximum number of characters) and restrictions on characters (“do not use commas”) are valuable information. Some hints about how to create a password that’s hard to guess would also be appropriate. It’s this last type of information that’s difficult to create—it requires deep knowledge of the subject matter rather than just an understanding of the product’s features and limitations.
- Integrate the documentation with user-generated content. Users should be able to search a company’s entire website—knowledge bases, forums, blogs, wikis, and documentation—for the information they need. It would be wise to ensure that the origin of the documentation is clearly labeled, so that users know whether information is official (authored by the product owner) or unofficial (contributed by end users).

Closing the Gap with XML

XML can help you address several of the bullet points in the preceding section:

- Accelerate document publishing. By automating formatting and publishing, XML-driven workflows can reduce or eliminate

production time.

- Provide content online. An XML-based workflow lets you deliver information online (in HTML or other formats) with commenting and other features.
- Integrate with user-generated content. XML can form the backbone of this integration. As we move into the new world of many-to-many publishing, official product information will become just one of many sources available to readers. Product owners who isolate their official information from the conversation going on about their product run the risk that their voice will not be heard at all. So here are my recommendations:

1. Move your product documentation into an XML-based workflow. XML will help you when you are ready to start integrating your documentation with other content sources. (Scott Abel likes the term “remixing” for this process.)
2. Provide your end users with avenues to contribute to the product documentation. You could, for example, open up commenting on your documentation as Adobe LiveDocs help has done (see www.adobe.com/support/documentation/ and http://help.adobe.com/en_US/InDesign/6.0/, for a specific example with InDesign).
3. Get involved with user-generated content as a curator—highlighting particularly useful contributions, pointing people to information in the product documentation that answers their questions, and moderating forums as necessary.
4. Combine product documentation with user contributions. Link from forum posts into the help, use

feedback to identify areas where help needs improvement, and take comments seriously.

5. Let users become contributors. Consider allowing end users to contribute to your documentation. (The WebWorks wiki, wiki.webworks.com, is one of the earliest examples of this approach in technical communication.)

There are obvious pitfalls, starting with liability and regulatory concerns, when you start to open up product documentation. In some industries, you may be constrained by safety concerns. However, consider the possibilities if you can harness the energy of your user community to help you improve the product documentation. Clay Shirky refers to the “cognitive surplus”—the free time that we collectively waste watching TV and that we could put to better use in “producing and sharing” information. We need to give our readers the opportunity to participate and see where it takes us.

Article archives for the XML Strategist and assorted white papers can be found at www.scriptorium.com/papers.html. This article is based in part on my “Web 2.0 and Truth” post. All blogs and podcasts referenced in this column were accessed 18 October 2008. 

SUGGESTED READING

Abel, Scott. “It’s in the Mix: The Next Generation of Open Source Publishing.” *The Content Wrangler* blog, www.thecontentwrangler.com/article/its_in_the_mix_the_next_generation_of_open_source_publishing/.

“Friend or Foe? The Role of Web 2.0 in Technical Communication” white paper, Scriptorium Publishing, www.scriptorium.com/whitepapers/web2/.

Gentle, Anne. *JustWriteClick* blog (see especially the social media category), <http://justwriteclick.com/>.

Johnson, Tom. “The Convergence of Web 2.0 with Help Documentation” podcast. *I’d Rather Be Writing* blog, www.idratherbewriting.com/2007/01/03/help-20-the-convergence-of-web-20-with-help-documentation/.

O’Keefe, Sarah. “Web 2.0 and Truth” post. Palimpsest blog, www.scriptorium.com/palimpsest/2008/06/web-20-and-truth.html.

Shirky, Clay. “Gin, Television, and Social Surplus.” *Here Comes Everybody* blog, www.shirky.com/hercomeseverybody/2008/04/looking-for-the-mouse.html.

Sarah O’Keefe (xmlstrategist@scriptorium.com) is founder and president of Scriptorium Publishing Services, Inc. (www.scriptorium.com) based in Research Triangle Park, North Carolina. The company is focused on implementing tools and processes to optimize publishing workflows. Services include developing and deploying XML-based structured authoring environments, configuring authoring and publishing tools, and providing technical training. Her publications include Publishing Fundamentals: FrameMaker 7, The WebWorks Publisher Cookbook, Technical Writing 101, FrameMaker for Dummies, and numerous white papers.