



Editor: James Hendler
Rensselaer Polytechnic Institute
hendler@cs.rpi.edu

Reinventing Academic Publishing—Part I

Intelligent Readers,

Although quoting yourself is generally considered tacky, I've been involved in several recent activities and discussions I'd like to share with you. These largely arose from "Publishing on the Semantic Web," a column that Tim Berners-Lee and I coauthored in *Nature* back in 2001 (www.nature.com/nature/debates/e-access/Articles/bernerslee.htm). In that column, one of a series of opinion pieces about academic publishing's future, we discussed the Semantic Web's potential impact. We ended with this somewhat brash statement:

The semantic web will provide unifying underlying technologies to allow these concepts to be progressively linked into a universal web of knowledge, and will therefore help to break down the walls erected by lack of communication, and allow researchers to find and understand products from other scientific disciplines. The very notion of a journal of medicine separate from a journal of bioinformatics, separate from the writings of physicists, chemists, psychologists and even kindergarten teachers, will someday become as out of date as the print journal is becoming to our graduate students.

In the past year, I've found that this quote is resonating more and more and that some of the big players in academic publishing are starting to think along these lines (perhaps not yet the kindergarten-teacher part, but that will come). Not all of them are considering using Semantic Web technology; some are inspired instead by Web 2.0's community-oriented features. However, the notion of breaking down the lines between traditional disciplines and reaching out to audiences beyond the academic bench scientist is becoming an important "meme" in academic publishing.

More than just technology

One leader in this area has been *Nature* itself, which launched the Nature Network (<http://network.nature.com>), a social-networking and blogging site aimed primarily at scientists. Users can create a social network, share forums and blogs, and use tags to create semantics. Given *Nature's* early interest in RDF and Semantic Web technologies, I hope some ontologies might eventually be added to improve linking between tags and to help provide a mechanism for cross-disciplinary searching. How successful the network has been depends on whom you ask, but it's a clear indicator of things to come when a major force in scientific publishing is exploring how to use Web technologies to enhance scientific communication.

Lately, I've been hearing from other publishers and magazines that they're also considering doing more to enhance their online sites with community-oriented features. They have been motivated by *Nature's* lead, by users' increasing reluctance to pay for hardcopy articles when so much is free online, and by the increasing facility that young scientists, the desirable demographic for the industry's future, have with Web technologies. Some of these publishers are just "jumping in"—throwing up sites to see what happens; others are taking a more cautious approach. Such caution is warranted—the success of a few Web 2.0 sites is causing too many people to think it's a silver bullet, without considering the nature of this technology's success, and failure.

It's becoming clear that making a successful community-oriented Web site requires more than just the technology. As a cautionary anecdote, consider the success of Wikipedia and the lack of success of a number of other sites that have tried to use wikis similarly. Many of these sites used technology virtually identical to the MediaWiki.org code from which Wikipedia is built, but have had nothing near the original's success. Some studies have begun to explore why this is the case, but anecdotally the key appears to be somehow in the social

We'd Like to Hear from You

Letters to the Editor:

Send letters, including a reference to the article in question, to intelligent@computer.org. Letters will be edited for clarity and length.

Articles:

If you're interested in submitting an article for publication, see our author guidelines at www.computer.org/intelligent/author.htm.

structures the successful sites have created. Similar stories arise in exploring Flickr, YouTube, and other Web 2.0 marvels. While these sites have taken off, others have languished owing to misunderstanding the mismatch between the technology and the community they want to reach.

Overcoming resistance

Which brings us back to academic publishing. As publishers try to promote new models of communication among scientists, with an eye toward finding some new role in the process, they need to respect the way science works. Although this, like many other things, might be changing owing to the Web's impact, some natural points of resistance must be overcome before new, more community-oriented, interdisciplinary scientific sites succeed. While scientists have gloried in the Web's disruptive effect on publishers and libraries, with many fields strongly pushing open publication models, we're much more resistant to letting it disrupt the practice of our disciplines.

At the Science Foo Camp (tagged on many blogging sites as "scifoo") held at Google in August 2007, several sessions dealt with academic publishing's future. The topics included open-source publishing; publishing "pre-review" (or with community reviewing of some sort); and the use of blogs, wikis, and other new technologies to enhance scientific communication.

However, motivated by comments arising the first day, the second morning featured the session "Culture of Fear: Scientific Communication and Young Scientists." This session, led by postdoctoral researchers Alex Palazzo (Harvard) and Andrew Walkingshaw (Cambridge), explored issues that those starting out in scientific fields face when using these new technologies. The job market for scientific positions, especially in academia, is tight. So how do a team of scientists, sharing partial results pre-publication, assign credit? Authorship blurs when small amounts of information, which might contain key insights into making processes successful, are publically shared. How does a blogger get credit for the information that leads to an eventual publication at a competing lab?

Another theme of the session was peer review's role in scientific fields. Although some pseudoscientists have claimed that we use peer review to keep their brilliant insights out of our precious literature, most scientists truly appreciate the filter that peer

review provides. The high standards publications maintain are a useful way to ensure that ideas are well argued and strongly evaluated before being published. On the other hand, some feeling has always existed in the community that, especially with respect to funding, the peer review process might be overly constraining and considerably delay new ideas from coming to the fore. The difficulty and delays that promising young scientists must face in the current system also affect hiring and promotion. These factors have motivated many in the community to discuss new mechanisms, based on emerging Web technologies, that let us communicate more ideas more quickly. For example, one model involves publishing after a minimal peer review and then creating some sort of postpublication metrics as to the paper's value. Some online journals are al-

As publishers try to promote
new models of communication
among scientists, with an eye
toward finding some new role
in the process, they need to
respect the way science works.

ready exploring this model. However, as Alex put it in his blog (http://scienceblogs.com/transcript/2007/08/scifoo_day_3_well_that_was_yes.php),

Until the scientific establishment reaches a consensus as to whether these post-publication metrics are indeed useful for determining the credentials of a scientist in the shorter term (<2 years post-publication) it is unlikely that any scientists would risk publishing their findings in a minimally peer-reviewed journal.

What can we do?

So, we arrive at the crux of the issue facing many of us, whether we're the editor in chief of a magazine such as *IEEE Intelligent Systems* or the head of a major press or publishing house. How do we embrace the new technology and encourage more of the sharing that Tim and I were calling for, without causing career risk to the very peo-

ple to whom the technology is most familiar—the younger scientists? If we don't think through the social issues of usage, the technologies alone won't have any significant impact and will go largely unused.

One option—and I'd like to see more effort in this area—is for innovation to come "from the top." Eventually, as these young scientists become the leaders of our fields, they will bring these new technologies with them. But with the world in its current shape, needing the help of scientists for our very existence, we can't afford to wait that long. Rather, we need to find ways to bring more senior scientists into contact with the positive side of these technologies.

In computer science, where the barrier has been lower than in some fields for senior people to learn to use new computer technologies, we've seen some of this happen. For example, Tim Finin has been instrumental in bringing bloggers to the AAAI conference, which has caused others to read, and in some cases create, blog content. Tim Berners-Lee's first blog was greeted by stories in newspapers around the world, making it that much easier for the rest of us to legitimize the time we spend sharing our thoughts in this lightweight way. When young scientists see their field's leaders embracing new technologies, it's that much easier for them to demonstrate to the rest of us, without fear of retribution, what these technologies can do.

It's time for us as computer scientists to take a leading role in creating innovation in this area. Some ideas are simple—for example, providing overlay journals that link existing Web publications, thus increasing the visibility (and therefore impact) of research that cuts across fields. Others might require more work, such as exploring how we can easily embed semantic markup into authoring tools and return some value (for example, automatic reference suggestions) through user-extensible ontologies. In my next column, I'll discuss current ideas regarding new technologies for academic communication that we as a field might be able to help bring into being, and some of the obstacles thereto. I look forward to hearing your thoughts on the subject. ■

James Herdler