John Dove

✪Consultant at Paloma & Associates

Overcoming Inertia in Green Open Access

Jun 24, 2015

There's a problem in Green Open Access.

Great strides have been made to pursue the objective of having scholars worldwide have unfettered access to the body of work that represents the state of knowledge in each field of inquiry. There are now a good number of peer-reviewed journals which are open access from day one. A large number of journals are now offered with a hybrid model where funders or authors themselves can sponsor the open access of an article right at publication time. Almost all of the top scholarly journal publishers worldwide have open access journals of their own. And almost all publishers acknowledge that the lead authors of an article can, by right, (sometimes after an embargo period) self-archive a version of their article in an institutional archive that is open access.

Some of the most important funding organizations in the world have mandates that the research which they fund must be reported on in open access journals (again, sometimes after an embargo). Now close to 100 universities worldwide have adopted open access policies strongly encouraging scholars at their universities to post the results of their scholarly work in an institutional repository.

The success of the Open Access efforts over the past 15 years can be seen by looking at the list of members of the Open Access Scholarly Publishers Association. http://oaspa.org/ membership/members/ It includes almost every one of the scholarly publishers in the world as well as the leading copyright rights management organization (CCC).

However, even with mandates, policies, and endorsement from publishers only a small percentage of scholarly journal articles worldwide are ending up open to all scholars to read. It's apparently the case that wherever there are mandates or policies in place a good rate of adoption of open access is achieved, but this represents only a small portion of all scholarly work and does not cover research that was done before the mandates or policies were in place.

I've heard from talking to a senior person at SPARC that for some subject areas the self-archiving rate in places where there is no mandate or policy is less than 15%.

Here's a little exercise which I've now done looking at research papers in a wide variety of disciplines. Look at the referenced sources in a recently published paper. Unless you are reading this paper at one of the few fully funded research libraries you will find that a majority of the referenced sources are unavailable to you. Open access is simply not there. If you talk to any scholar or would-be scholar in any discipline about how they go about a careful reading of an important paper in their field you will find that a central part of the reader's experience is to browse through the full-text of several of the referenced sources. Very often these are papers written well before a mandate or policy on Open Access was in place in their discipline. If the reader is at a less than fully funded college or university, or is reading the article from a less

developed country, an attempt to do a detailed study of an article is a throw-back to days before computers. Lots of the referenced sources will have to be obtained by inter-library loan or not at all. Your ability to participate in the scholarly inquiries of your field is highly constrained.

My brother is a researcher at University of Wisconsin with a long published record in molecular genetics. His research is all funded by NIH so he is required to publish results in ways which will achieve open access. But when I look at articles that he's published in the last two years, the majority of sources he references in his papers pre-date the mandates and policies. I've looked as well at the dozen papers which are published by PLOS (the leading purely Open Access publisher) which reference my brother's work and the majority are articles which he wrote before there were policies or mandates and some are still behind paywalls.

So, how can the vast majority of scholarly articles that represent the basis for scholarly work make their way into an accessible place so that the mission of Open Access can be accomplished? It does not surprise me that there is inertia among scholars to exercise their rights to self-archive. Scholars are busy. It generally takes an active step by authors to self-archive. Of course, publishers could open up access after some period of time and same may do so, but there's no obvious self-interest among commercial publishers that would drive them to do this. And there is a lot of misinformation about open access.

Publishers are missing a trick

I think there's a case to be made that journal publishers may be missing a trick. There is a point in time when a publisher's self-interest in the quality of their about-to-be-published work would be well-served by encouraging authors of referenced sources to self-archive their past articles. This is also a moment in time at which the authors of referenced sources are also missing a trick but are unaware of it.

Imagine the publisher of an article about to be published. They could examine the soon-to-be-published article and take note of the referenced sources in that article. Which of them are originally published in an open access journal or have had a version of the article archived in an institutional repository? The utility of this about-to-be-published article is clearly enhanced by having as many referenced sources as possible be eventually made open. That way learners can fully absorb the import of this article as they can browse the underlying research on which the new results are based. [Nothing in what I propose should ever imply that selection of which articles to reference should be based on whether they are open or not. Quality and recognition of prior work needs to trump openness. That seems to me to be a foundational principle of academic integrity.]

So I am proposing that the publisher ping the authors of referenced sources which are not openly accessible with some version of the following message:

• Good news! We are about to publish a peer-reviewed article that cites your article such-and-such. This is a tribute to your good work.
• We notice that your article is not openly accessible, yet could be.
• Here's how to do it: xxx-yyy-zzz.
• Did you know that if your article is openly accessible not only will it be read more times but it is also more likely to be cited in the future? [I understand that empirical research backs up this assertion.]
• If you have any questions let us know
• Keep up the good work

In cases where there are multiple authors of a referenced source I think the publisher should ping all of the authors even though in some cases it's only the lead author who is recognized to have the self-archive rights. This is because the lead author may already have tenure and therefore may not be as strongly self-interested in further recognition. But her co-authors are very likely to be her graduate students and post-docs and protégés. Perhaps some social pressure will help nudge the lead author to exercise their self-archival rights.

As one senior executive of a large scholarly press pointed out to me in discussing this idea, communicating to authors of referenced sources is an opportunity to strengthen their brand with important authors in the field. So the self-interest for the publisher in doing this encouragement of opening the referenced sources is more than just to improve the quality of the new article; it's also about establishing a presence with potential future contributors to their journals.

Publishers who are purely open access publishers have another motivation to undertake this pinging process. They are not just scholarly publishers--some also have a mission to actively advocate for open access. Pinging these authors of referenced sources is a great opportunity to educate broadly among scholars in the disciplines that they publish in. And answering questions that come back from these authors can be an important measure of how well scholars understand the plethora of issues that continually arise in the industry.

Another interesting effect of this pinging proposal is that pinging will tend to gang up on the most important articles since they will be the ones that are more often cited.

What's wrong with this proposal?

I always remember the advice of an old friend, Gerry Weinberg, from his book *Are Your Lights On?*, "If you can't think of at least three things wrong with your proposal you probably don't understand the problem." So here are some things potentially wrong, many of them raised by people I've shopped this idea with recently:

• While it is clear that articles with referenced sources that can be clicked through to access of the full text are better, is the cost worth the benefit?

• If not, are there enough clear ways to automate much of the process so that the costs go down?

• Is a publisher likely to have access to the contact information for a sufficient number of the referenced sources to make this a useful effort?

• Are changes needed to citation standards, link-resolvers, DOIs or other things that effect a user's ability to locate referenced sources which may reside in different places?

• What about referenced sources where the author is no longer living or accessible?

Some other things that could be done

• Everything said here about publishing a journal article could equally be said about a couple of other content types important to academic publishing, namely: annotated bibliographies and curated lists of resources. Again, inclusion of a referenced source that is open is a better user experience of that bibliography and is also a moment of good news to the author that their work is now being recognized. A bibliography where most or all of the referenced sources are available is clearly of greater use than one which simply identifies sources which are inaccessible.

• A couple of tools have recently been developed and made available to everyone that assess openness. One, the Open Article Gauge (http://howopenisit.org/about) reports back to inquirers what restrictions exist around various uses of scholarly articles. This was developed

as a joint effort by PLOS, SPARC, and OASPA. Another is the Open Access Button, a browser plug-in which helps users when confronted with a paywall to find an open access version or e-mail the author to request a copy. Perhaps these tools include some of the underlying software on which to build a tool that assesses accessibility of a reference list.

• People tell me that there is not yet an effective aggregator of Open Access journals. As publishers step forward to fill this gap one could imagine a variety of features in an aggregated product offering that could shed light on how many referenced sources are open, or could facilitate users being able to get to alternative versions of articles that are more open than others.

I think that the combination of these measures could definitely "move the needle" in addressing misunderstandings among scholars about open access and begin to overcome the inertia that otherwise holds back the full realization of the benefits of open access to the world's scientific and scholarly enterprise. Keeping an eye on the total utility of a journal article to the least resourced scholar (other than access to the Internet) I think we can deploy continuous improvement until the full benefits of open access for scholarly journals is achieved.

Feedback so far

I've now shopped around this basic idea to a wide variety of people including senior management of a couple significant academic publishers, scholars (among friends and family) in a wide variety of disciplines including STEM, social sciences, and humanities, well recognized experts in Open Access, Alt Metrics, members of NISO, librarians, and open access publishers. Everyone so far has found the idea intriguing and worth further examination. One put it this way, "I'm not aware of publishers doing this today and I think it's definitely something they should try."
I would like to single out Peter Suber's Book, Open Access published by MIT Press https://mitpress.mit.edu/books/open-access and now available open access) which was especially helpful to me to clarify my ideas about this suggestion.

I welcome feedback on this idea and even offers to collaborate on a more detailed rendition of this proposal that might fill in some of the gaps in this proposal or correct errors in my assertions.