

Cites & Insights

Crawford at Large

Libraries • Policy • Technology • Media

Sponsored by YBP Library Services

Volume 9, Number 12: November 2009

ISSN 1534-0937

Walt Crawford

Library Access to Scholarship

It's been seven months since the last LIBRARY ACCESS TO SCHOLARSHIP—and that edition barely touched issues of open access, focusing solely on the (possible? desirable?) death of journals. Realistically, it's been a full year since *Cites & Insights* has had any discussion of open access and the controversies and updates swirling around the concept.

It's time to catch up—and maybe it's time to give up. More on that at the end of this jumbled essay, consisting of excerpts from dozens of items on aspects of OA, with my own comments as appropriate. How many items? I began with more than 90, winnowed that down to 69, and probably ended up with around 60.

Items are arranged by general topic and, typically, chronologically within a topic. Most *italicized* sub-headings are blog post or article titles. Almost everything is from the past year, sometimes going back earlier in 2008. Don't expect neat narratives in each section; in many cases, these are items I thought worth noting even if they don't connect very well. As is general practice in *C&I*, I use sentence capitalization for most blog names—and only provide URLs when they're not readily discoverable.

Mandates, Policies and Compacts

Harvard has an OA mandate. NIH now has an OA mandate, upgraded from a request that wasn't being heeded. Many would like to see a broader Federal government OA mandate for government-sponsored research. These are mostly relatively smaller items in that overall landscape.

Freedom, Google-juice and institutional mandates

Andy Powell (of eduserv) posted this February 10, 2009 at *eFoundations*. He discusses a thread on a forum asserting that, in Germany, freedom of research *forbids* institutional mandates—it wouldn't be legal for

a German university to mandate that faculty deposit research papers into an institutional repository.

One response, from Bernard Rentier, said:

No researcher would complain (and consider it an infringement upon his/ her academic freedom to publish) if we mandated them to deposit reprints at the local library. It would be just another duty like they have many others. It would not be terribly useful, needless to say, but it would not cause an uproar. Qualitatively, nothing changes. Quantitatively, readership explodes.

To which Powell says the web isn't like a library, so the analogy isn't a good one. After trashing OAI-PMH as irrelevant ("If we ignore the rarefied, *and largely useless*, world of resource discovery based on the OAI-PMH") Powell asserts that within the "real world,"

[t]here is a direct and negative impact of mandating a particular place of deposit. For every additional place that a research paper surfaces on the Web there is a likely reduction in the Google-juice associated with each instance caused by an overall diffusion of inbound links.

So, he says, if a scholar prefers to "surface their paper" somewhere else on the web, but is also forced to put it in the IR because of a mandate, they will "probably see a negative impact on the Google-juice associated with their chosen location." He also objects to thinking of a repository as "part of a library" rather than as "a true part of the fabric of the web." In summary, he regards repository mandates as "a disservice...to the long term future of open access."

Right off the bat, I wonder how the desideratum for scholarly success has shifted from publication, to impact, to readership—and now to *Google-juice*? Is it *really* likely that a paper existing on two sites will be read less frequently than the same paper existing on only one—or is it likely that the paper will be read more frequently but in more distributed fashion?

In any case, Powell recognizes he was partly wrong (the first version of the post appeared a day earlier): Google joins multiple instances of a research paper in its indexing, at least in many cases. Which “somewhat” weakens his argument for “dispersion of Google juice.” So he throws in other arguments—e.g., that W3C argues against having different URIs for the same resource. Still, it seems Powell thinks Google PageRank is *important*—and that maximizing GPR is a valid scholarly enterprise. This is, to me, a rather startling shift in the aims of the scholarly enterprise. The comments are interesting, including one that pointedly says, “I don’t get promoted on PageRank.”

On the way to a mandate

“An interesting thing happened on the way to universal open-access mandates...”

That’s how Dorothea Salo begins this February 18, 2009 post on the dearly departed *Caveat lector* (cavlec.yarinareth.net, because there’s more than one blog with that title). She notes the sad history of university *resolutions* in favor of OA, beginning in 2007:

As a symbolic action, this was lovely and welcome. As an instrument for making more journal articles open-access, it was perfectly useless. Not only was there no actual requirement that faculty retain rights (which of course tends to mean they won’t do it), there was no reporting or recording involved anywhere (that I know of, anyway).

As a result, most resolutions had little or no effect, at least in terms of increasing access to the articles. So Harvard took a different route: A mandate to *deposit* copies of all articles.

Personally, I’m still not sure the average faculty member covered under that resolution is completely cognizant of the [the deposit requirement], and I do still believe things could get ugly once they figure it out. I hope not...

But she’s mostly interested in a Boston University faculty move—an open access *plan* that isn’t a mandate or resolution:

Instead, it is a fascinating middle-ground. It mentions gold as well as green OA. It mentions building a faculty publications database, not just an IR; this is important because like it or not, faculty publications databases have real-world uses for faculty and administrators that IRs simply don’t. It takes on tenure and promotion practices straightforwardly.

It is, in short, a start toward a university-wide open-access strategy. That’s fascinating, and to the best of my knowledge, completely novel. The breadth of the conversation is certainly a vast improvement over the library starting an IR all by itself that it then doesn’t

promote or work to fill. It’s also an improvement over putting all the local open-access eggs in one basket, whether that basket is an IR or an author’s addendum or a gold-fee fund...

Noted here as part of a mandate thread, but note that last sentence when we get to philosophical discussions (or knockdown drag-out Green vs. Gold battles) later.

How we’re walking the OA walk

From *ACRLog*, on May 17, 2009, by Barbara Fister at Gustavus Adolphus College—a story about a smaller college and the *library* faculty at that college.

The good news about open access keeps coming. Here at *ACRLog*, we’ve followed the trend since Harvard’s Arts and Sciences faculty adopted an open access resolution. Boston University and MIT have made similar resolutions... Recently Oregon State University librarians adopted an open access mandate, followed quickly by the University of Oregon...

Well, at our last librarian’s meeting we adopted our own Open Access Pledge. It’s not as sophisticated as the ones that have been making news. We are a small library, with only six librarians, and we haven’t had the time or money to start up an institutional repository. We also, quite frankly, don’t have a terribly sophisticated grasp of all the OA arguments, the copyright issues, and the color choices. (Green? Gold? What about mauve?)...

...Our pledge is simply to make every effort to ensure that our scholarship is freely available online, either because the publisher posts its content online (as does *Inside Higher Ed* or *Library Journal*), it’s a truly OA journal, or because the publication agreement allows self-archiving, which most credible library publications do. We also pledge to do the work of self-archiving, which really isn’t a lot of trouble for librarians who are tweaking the web daily. It mystifies me that so few librarians can be bothered.

This wasn’t a simple decision. Half of the department is on the tenure track. Their continuing employment depends on establishing professional credibility through publication. But we feel strongly that this is the right thing to do, and that taking these simple steps won’t damage the ability of our emerging scholars to thrive...

The comments are generally positive. When one asks where (without an IR) library faculty plan to post their articles, Fister mentions server space—but fails to note that there is already at least one topical repository, E-LIS, with more than 9,500 articles as of late September 2009.

Here’s the pledge itself:

The Gustavus library faculty believes that open access to scholarship is critical for scholarly communication

and for the future of libraries. For that reason we pledge to make our own research freely available whenever possible by seeking publishers that have either adopted open access policies, publish contents online without restriction, and/or allow authors to self-archive their publications on the web. We pledge to link to and/or self-archive our publications to make them freely accessible.

Librarians may submit their work to a publication that does not follow open access principles and will not allow self archiving only if it is clearly the best or only option for publication; however, librarians will actively seek out publishers that allow them to make their research available freely online and, when necessary, will negotiate with publishers to improve publication agreements.

Open Access mandates: Judging success

That's Richard Poynder's title for a May 23, 2009 article at *Open and shut?* He notes the spread of institutional mandates and wonders about compliance.

[W]hat level of compliance can we expect from these mandates? After all, a mandate is only as good as the compliance rate it achieves. And how do we judge success so far as compliance is concerned anyway?

Poynder notes (and links to) analyses by Arthur Sales within Australia indicating a 70% compliance rate within two years of a mandate—and the clear effects of NIH's move from request (5% compliance) to mandate (around half of NIH's target in the first year or so).

A mandatory self-archiving policy at the University of Stirling has not yet achieved anything like ubiquitous compliance—but deposit rates in the University's IR have jumped from 20 per month to 120 per month.

Poynder poses several questions deserving of medium-term study by impartial observers:

Can we expect the NIH compliance rate to match the levels reported by Sale in Australia by next April (i.e. two years after it became mandatory)?

Can we expect the surge of new mandates to achieve the same levels of compliance reported by Sale?

Is there any significance in the fact that many of the new mandates are being introduced by library faculties, and can we expect that to affect compliance rates?

Will the fact that many of the new mandates are self-imposed affect compliance rates? (Will it make them appear more voluntary than mandatory)?

Will the fact that many of the new mandates include opt-outs affect compliance rates? (Will that make them appear more voluntary than mandatory)?

What is full compliance so far as a self-archiving mandate is concerned? (Is Sale's 70% level the objective, or should the research community be aiming higher?)

What other questions should we be asking, particularly when trying to judge the success of a mandate?

All excellent questions, none with simple answers.

Are the Harvard open-access policies unfair to publishers?

That's the provocative title of Stuart Shieber's June 9, 2009 post in his excellent new blog, *The occasional pamphlet* (blogs.law.harvard.edu/pamphlet/). Getting past a one-word two-letter answer that might be sufficient, some excerpts from Shieber's discussion:

Recently, the representative of a major scientific journal publisher expressed to me the sentiment that the position that Harvard faculty have taken through our open-access policies—setting the default for rights retention to retain rights by default rather than to eschew rights by default—is in some sense unfair to subscription-based journals that require embargoes, that we are favoring one scholarly publishing business model over another and setting up an unlevel playing field...

Shieber summarizes the policy and its waiver provisions, noting that the publisher in question felt the policy was unfair because OA publishers didn't need the author to obtain a waiver while toll-access publishers did.

Of course, as a private company, Harvard is well within its rights to set up its policies to favor whatever it wants, even if third parties are disadvantaged... But I for one am not interested in unfairly advantaging one business model for scholarly publishing over another, and in any case, the argument that the Harvard policy does so is fallacious.

Before describing why, let me mention again that there is a different issue that I will not address here, that immediate open access of author's manuscripts endangers the subscription-based business model. That is a separate issue, with a set of standard arguments and counterarguments...

This other argument is different and seems to rely on an assumption that the Harvard policy is an undue intervention, whereas without the policy, there is no intervention. It seems to go something like this: In designing a system of rights allocation between author and publisher, the approach in which all rights are transferred to the publisher with the publisher selecting some rights to provide back to the author is the privileged position, and any other arrangement involves some kind of intervention. But what makes that the privileged (nonintervening) position? Why isn't the privileged position an approach in which all rights are transferred to the publisher subject to a nonexclusive limited noncommercial license to the author, with a publisher's requirement for exclusive transfer of all copyright being the intervention? It is true that one was the status quo ante for some time,

but that is a historical contingency, and gives it no privileged position as consisting of the appropriate default position. There is in fact no privileged position one way or the other. All choices of design involve intervention. There are just a large variety of possibilities, which the actions of various stakeholders choose among.

I'd go further: Why isn't the privileged default position for scholarly articles the same as it is for most (paid) nonscholarly articles in most periodicals—where the *only* transfer of rights is for first serial publication, with the writer retaining *all* other rights? That might also hinder immediate green OA (or it might not), but it has an honorable history.

Another way of thinking about the issue is this: The faculty must make a decision about what portion of their rights they retain by default: none, a few, a lot, this subset, that subset. A decision is inevitable; there is no option of making no decision. The status quo ante made a decision, namely that no rights were retained by default. The Harvard policy makes a different decision, that a nonexclusive limited license is retained by default; some other policy might state that exclusive rights are retained by default. But one way or another, a decision must be made. The faculty are obviously inclined to make decisions that benefit them as they perceive it, and have done so. The change is not from making no decision to making a decision, it is from making one decision to making another that the faculty perceives as preferable...

I appreciate that some publishers may not like the waiver process, and I do not begrudge them that view or think it is unreasonable. But preferring X to Y doesn't mean that Y is unfair.

There's more to the original post (and, as with other Shieber essays, it's worth reading). I've always been taken with the concept that publishers might *decline* Harvard submissions because of the mandate—a policy decision that would, I believe, be somewhere between idiotic and suicidal. (I have no evidence that this has ever happened, nor do I think it likely.)

Quantifying the value of peer review

If most OA bloggers are believers (or steadfast opponents) and I'm an independent, T. Scott Plutchak revels in the title "OA heretic." In this June 10, 2009 post at *T. Scott*, he advances an interesting argument against the NIH mandate. Excerpts:

...Despite all of the blather and charges and countercharges we still haven't really sorted out the relationship between the value of peer review, the role of the publishers, and the impact on the public good.

Explicit in the NIH policy is that peer review has substantial value—so much so, that NIH does not want

any manuscripts deposited that have not gone through a rigorous peer review process and gotten the stamp of approval from a recognized peer review authority—i.e., a publisher. In developing the policy, NIH could have come up with their own vetting mechanism, but instead they quite sensibly chose to rely on the experts in managing peer review. (And don't be fooled by the oft repeated truism that "peer review is all done by volunteers anyway." If it were that simple, why wouldn't NIH just set up their own peer review system?)

But, T. Scott, you're not denying the *truth* of the truism (outside of a few special cases). Maybe NIH has not set up their own peer review system because (a) such systems already exist and (b) if they did, they would *immediately* be charged with unfair government competition. I call straw man.

But here's where it gets sticky. In "the old days" (when everybody understood what the rules were), publishers gained control of copyright in exchange for managing the peer review process. They were then entitled to use that control to develop revenue streams that would compensate them for the value that they were adding to the system. Copyright gave them control of the distribution of the work to which they had added value. Under the terms of the NIH policy publishers are expected to give up that control. And it irks them.

So it's a simple quid pro quo: Peer review *management* (not peer review—publishers don't do that) in trade for *all* rights? I suspect most scholars never saw it that way. And the NIH policy does *not* take away control. It provides one exception for one class of papers. An exception, by the way, that would be automatic for government-funded research done in government facilities.

Publishers add value in many ways... All of the elements that visually distinguish the final published article from the author's manuscript version can be useful and worthwhile. The public access policy says that the publisher is still free to get compensated for all of those things, as if they were the totality of the value that the publisher adds.

But the publisher has already lost control of what is most important—the imprimatur that says that this article has gone through the peer review process and has been accepted as worthy. Since that process doesn't necessarily result in significant changes of expression, it's not part of the copyrightable package, but it adheres to the author's manuscript version every bit as much as it does to the final published version.

The *most important* aspect of journal publication is "this article passed peer review"? Maybe.

It is argued that this is not an unfair "taking" since the publisher has the right to refuse to grant the license

that allows the author to deposit with Pubmed Central. Puh-leeze! This is, no doubt, technically and legalistically true. But since when is a choice between complying with a policy and going out of business a real choice? “Dear publisher—we respectfully ask that, for the benefit of the common good, you give up control of the most significant element of value that you add to the scholarly communication process. We don’t actually have any way of compensating you for that, so you are perfectly free to refuse to do so—in which case, you will, of course, be put out of business since you will no longer receive the manuscripts that are your bread and butter, but them’s the breaks. Good luck.”

Again, remember that publishers only *manage* peer review—but, according to Scott, that’s the “most significant element of value” that they add. What he’s asserting here is that government-funded research is the “bread and butter” of medical publishers, but that it’s *unfair to publishers* for taxpayers to assert any rights in what they’ve paid for.

...While everyone involved in the debate agrees on the fundamental importance of peer review, and everyone seems to agree that it ought to be managed by the traditional peer review authorities, the pro-policy advocates have not directly confronted the fact that the policy does require publishers to give up control of something of great value without providing anything in return. Wouldn’t it be more constructive to acknowledge this fact and try to develop policies that account for it in some way, rather than dodging behind legalisms and hyperbolic rhetoric?

Really, I’m just lookin’ for a little intellectual good faith here.

Am I being too hard on T. Scott? Perhaps, but I don’t think so. I think his claim (that *managing* peer review is the *primary* value added by publishers and is so important that it *alone* justifies seizing all rights in publicly-funded research) is extreme. Or, say, wrong.

The first comment suggests that publisher management of peer review might not be such a big deal (and that, down the line, NIH indeed might set up a peer review process if it’s needed). The third, from another interesting recent voice in OA commentary (Bill Hooker), suggests that *everyone* should be entitled to be paid for the value they add:

My standard rate for peer review is \$25/hr. For OA journals, I will continue to review for free.

I would comment that \$25/hour is awfully cheap...

The OA policies of June

This brief item by Peter Suber appears in the July 2, 2009 *SPARC Open Access Newsletter*. In full (excluding links):

By my count, June 2009 was the richest month ever for OA policies. We saw new OA mandates adopted at the University of Bergen, the Harvard Graduate School of Education, and University College London. The University of Kansas and the nation of Lithuania adopted near-mandates. Older mandates were revealed at the University of Geneva, Roehampton University, and the US Institute of Education Sciences. An older mandate took effect at the National Cancer Institute of Canada, an older non-mandate was strengthened to a mandate at the Canadian Breast Cancer Research Alliance, and implementation began for an older mandate in Ukraine. The largest proposed mandate ever, the Federal Research Public Access Act, was re-introduced in the US Senate. Not counting calls for mandates and other policy recommendations, that comes to 12 developments in six countries (Canada, Lithuania, Norway, Switzerland, UK, Ukraine, and the US)—in one month alone.

For comparison, the second richest month was October 2006, when we saw six adopted OA mandates, two proposed mandates, two adopted near-mandates, and one adopted mandate limited to data, or altogether 11 actions in five countries (Austria, Canada, China, the UK, and the US).

Coming in third was January 2008, when we saw three new mandates, two start dates for previous mandates, one public disclosure of an older mandate, one tightening of an older mandate, and three near-mandates, or altogether 10 developments in eight countries (Australia, Canada, EU, Italy, Japan, Russia, Spain, Switzerland, and the US).

My only comments: Actions aren’t necessarily mandates—and mandates don’t necessarily spell success, although they usually improve levels of deposit. (This assumes that deposit constitutes success, which is another question—one I don’t intend to discuss.) Still, it’s evidence of momentum.

Maybe this last item is even more evidence, albeit in a very different direction.

A compact for open-access publication

That’s the title on a September 14, 2009 news item from Harvard University Library (hul.harvard.edu/news/2009_0914_compact.html). Since it’s a press release, I’m offering it in full, with notes following:

Five of the nation’s premier institutions of higher learning—Cornell, Dartmouth, Harvard, the Massachusetts Institute of Technology, and the University of California, Berkeley—today announced their joint commitment to a compact for open-access publication.

Open-access scholarly journals have arisen as an alternative to traditional publications that are founded on subscription and/or licensing fees. Open-access journals make their articles available freely to anyone,

while providing the same services common to all scholarly journals, such as management of the peer-review process, filtering, production, and distribution.

According to Thomas C. Leonard, university librarian at UC Berkeley, “Publishers and researchers know that it has never been easier to share the best work they produce with the world. But they also know that their traditional business model is creating new walls around discoveries. Universities can really help take down these walls and the open-access compact is a highly significant tool for the job.”

The economic downturn underscores the significance of open-access publications. With library resources strained by budget cuts, subscription and licensing fees for journals have come under increasing scrutiny, and alternative means for providing access to vital intellectual content are identified. Open-access journals provide a natural alternative.

As Dartmouth Provost Barry P. Scherr sees it, “Supporting open-access publishing is an important step in increasing readership of Dartmouth research and, ultimately, the impact of our research on the world.”

Since open-access journals do not charge subscription or other access fees, they must cover their operating expenses through other sources, including subventions, in-kind support, or, in a sizable minority of cases, processing fees paid by or on behalf of authors for submission to or publication in the journal. While academic research institutions support traditional journals by paying their subscription fees, no analogous means of support has existed to underwrite the growing roster of fee-based open-access journals.

Stuart Shieber, Harvard’s James O. Welch, Jr. and Virginia B. Welch Professor of Computer Science and Director of the University’s Office for Scholarly Communication, is the author of the five-member compact. According to Shieber, “Universities and funding agencies ought to provide equitable support for open-access publishing by subsidizing the processing fees that faculty incur when contributing to open-access publications. Right now, these fees are relatively rare. But if the research community supports open-access publishing and it gains in importance as we believe that it will, those fees could aggregate substantially over time. The compact ensures that support is available to eliminate these processing fees as a disincentive to open-access publishing.”

The compact supports equity of the business models by committing each university to the timely establishment of durable mechanisms for underwriting reasonable publication fees for open-access journal articles written by its faculty for which other institutions would not be expected to provide funds.

Additional universities are encouraged to visit the compact web site and sign on.

Cornell Provost Kent Fuchs offers his perspective on participating in the compact. “As part of its social commitment as a research university,” Fuchs says, “Cornell strives to ensure that scholarly research results are as widely available as possible. The Compact for Open-Access Publishing Equity could increase access to scholarly literature while at the same time ensuring that the valuable services that publishers provide are supported.”

A full account of the motivation for the compact can be found in the article “Equity for Open-Access Journal Publishing,” published in the open-access journal *Public Library of Science Biology*.

“Supporting OA journals is an investment in a superior system of scholarly communication,” states Peter Suber of the Scholarly Publishing and Academic Resources Coalition (SPARC) in Washington, DC, and a fellow of Harvard Law School’s Berkman Center and Harvard University’s Office for Scholarly Communication. “Before this compact, a number of funding agencies and universities were willing to pay OA journal processing fees on behalf of their grantees and faculty. It’s significant that five major universities recognize the need to join the effort, extend fee subsidies to a wider range of publishing scholars, enlist other institutions, and start to catch up with their long practice of supporting traditional—or non-OA—journals.”

Summing up the compact, MIT Provost L. Rafael Reif observes, “The dissemination of research findings to the public is not merely the right of research universities: it is their obligation. Open-access publishing promises to put more research in more hands and in more places around the world. This is a good enough reason for universities to embrace the guiding principles of this compact.”

The Compact website is at www.oacompact.org/ and uses COPE as an acronym (for **C**ompact for **O**pen-**A**ccess **P**ublishing **E**quity). The compact itself is brief (160 words!), but includes 14 hyperlinks that expand it considerably. Based on some of those links, I’m guessing a fair amount of additional fleshing-out is intended (e.g., it’s odd to have an FAQ page consisting entirely of a search box and introductory sentence).

This is about mandated institutional support for Gold OA, a very different approach to institutional mandates for Green OA. These are early days—the mandate appeared just a week before I began writing this article and the *day* I’d decided to stop collecting material—but this one should be unusually interesting to watch.

The Colors of OA

You could call this “philosophical discussions on the best approach to open access,” but at times that feels like call-

ing the Civil War a philosophical discussion on the best approach to labor relations. This time, it's not Blue vs. Gray, it's Gold vs. Green—even if the most thoughtful participants are almost always in the middle.

For those who need a quick reminder, Green OA is, basically, self-archiving or repository deposit of peer-reviewed articles—and Gold OA is, basically, journal articles that are freely readable from the journal itself. Gold OA means journals don't charge subscription fees *for electronic access to peer-reviewed articles*—although they can certainly charge for other things. Green OA means some peer-reviewed version of an article is freely available somewhere on the web.

Stevan “One Note” Harnad repeats, over and over and over and over again, at considerable length, that Only Green OA Matters (and it should be *his* version of Green OA, with repositories that magically cost almost nothing to run) and that any effort toward Gold OA is a diversion from The One True Path. These days, almost everybody else—except those trying to throw badly argued landmines into the whole procession—argues for some blend of the two. Here-with, a few brief incidents on the colors of OA.

Two-thirds full?

Dorothea Salo, *Caveat lector*, September 3, 2008. Salo refers (and links) to “a typically well-presented examination of journal prestige and how it plays out in toll-access/open-access faculty mindshare” by Peter Suber in the September 2, 2008 *SPARC Open Access Newsletter*. That article is worth reading; as usual, Suber writes well and thinks carefully—even if you may not always agree with him. One keen (in the non-slang sense) line: “I start with the rough notion that if journal quality is real excellence, then journal prestige is reputed excellence.”

Suber notes that universities have motivated faculty to publish in high-*prestige* journals, because such publications count more for promotion and tenure (and possibly for funding). But that may be the wrong motivation—and, given history, it's a tough obstacle for Gold OA advocates:

If journal prestige and journal quality can diverge, then universities and funders may be giving authors an incentive to aim only for prestige. If they wanted to create an incentive to put quality ahead of prestige, they haven't yet succeeded.

Suber quotes himself from a year earlier on the effect of rewarding publication in high-*prestige* journals:

By rewarding faculty who win a journal's imprimatur, mindful of the journal's prestige but heedless of its access policies, universities [and funders] shift bargaining power from authors to publishers of high-*prestige*

journals. They give publishers less incentive to modify their standard contracts and authors greater incentive to sign whatever publishers put in front of them.

As the discussion continues, most high-*prestige* journals are “toll access” (that is, aren't Gold OA)—and most authors will choose prestige over OA, not realizing that they *may not* have to choose (that is, many toll-access journals permit self-archiving, Green OA).

There's a *lot* more to Suber's discussion on how excellence can diverge from prestige and related issues—you need to read the article. But let's get back to Salo's commentary—which has mostly to do with one item in Suber's article stated as common knowledge: “About two-thirds of toll-access journals already give blanket permission for [self-archiving] and many of the others will give permission on request.”

Salo says:

This just isn't true, not unadorned, and I wish we'd stop waving it around. For it to be true, SHERPA/ROMEO (from whose database of publisher policies this datum is derived) would have to cover the entire toll-access journal universe. It doesn't. It doesn't even come close. Sure, it covers the behemoth toll-access publishers, but there are two problems with extrapolating from a set of data weighted heavily toward them: first, disciplinary coverage on SHERPA is extra-spotty in areas the behemoths can't profit from (notably the humanities); and second, I have to date seen zero evidence presented that the behemoths' policies are typical of non-behemoth publishers. I don't think they are, myself, though I'm willing to be wrong.

So Salo does a thought experiment, using personas she established earlier on (noted briefly in an earlier *C&I*)—the scholar “Dr. Troia” and repository librarian “Ulysses Acqua” both at “Achaea University”:

So let's see how this plays out with Dr. Troia and the repository librarian Ulysses Acqua. Dr. Troia comes to Ulysses with her CV and says “Let's put all this in the Achaea U repository!” Ulysses is of course thrilled, but because he is a responsible librarian, he explains carefully that someone has to check whether her publishers permit it. Chances are, he says honestly, that most do and a few won't.

Stepping out of frame for a moment—I've had faculty disengage *right then and there*, when they realize that they probably won't be able to get their complete publication record in. This is what they typically want—to use the repository as a publication-list proxy—and we often can't give it to them. Stubborn resistance on the part of certain repository-software packages to creating metadata records that do not include full-text files does not help this problem one little bit.

Let us say that Dr. Troia, though understandably disappointed, gives Ulysses the go-ahead to do the

checking for her. She naturally assumes that this is a service Ulysses offers; it doesn't even occur to her that she ought to do this herself. Ulysses asks diffidently if she will let him see the publication agreements she has signed. "I needed to keep those?" she asks.

This being the answer Ulysses expected (he is not so naïve as he was when he started his job), he smiles, reassures her he can manage without them, and makes a list of the journals she has published in. Let us suppose, somewhat unrealistically, that of those journals present in SHERPA/ROMEO, the distribution of permissive and forbidding publishers exactly mirrors the distribution of SHERPA/ROMEO as a whole: that is, two-thirds allow self-archiving in some form, while one-third forbid it.

Dr. Troia has published in some mildly obscure and boutique basketology journals, especially early in her career. Chasing down as much information as he can about them costs Ulysses about half a day's work. A couple have shut down entirely; being a pragmatic soul, Ulysses figures they're not terribly likely to sue him, and puts those articles on the "yes" pile. One more has transferred ownership twice since Dr. Troia published there. Ulysses groans, researches the old owner and the new (neither of whom is in SHERPA), and decides that on balance, this journal should probably be a "no." On one journal, Ulysses can find no information at all. His policy for such cases (discussed with his administration) is not to archive, though he is aware that some maverick repository managers do otherwise...

When he is done, Ulysses has about five-eighths of Dr. Troia's published articles in his "yes" pile. This isn't a bad or atypical result; depending on discipline, it can be anywhere from less than half to around 85%. He might be able to up the percentage a little bit if he were to negotiate with individual publishers, but he's tried that before, and the few positive results don't justify the time expenditure. He suspects the publishers would be more amenable if Dr. Troia were to contact them rather than a mere librarian, but he knows better than to ask it of her.

Now comes the *fun* part. Ulysses returns to Dr. Troia with a list of the articles that he needs her to provide preprints for, and another list of articles that he needs postprints for. At first Dr. Troia is confused; she can just download the publisher's version online, won't that do for him? Ulysses, his heart sinking, shows her a SHERPA page for one of the journals on the preprint-only list. He's fairly sure he knows what's coming, and sure enough, she says, "I couldn't find half these articles in draft if I tried." Because she's a nice person and likes Ulysses, she doesn't add, "and whyever would I bother trying? What a waste of time."

This leaves Ulysses with just the dead-journal articles and the articles archivable in publisher-PDF form....

Let's say one-quarter of Dr. Troia's CV, and that's fairly generous.

Ulysses isn't thrilled with that result. More importantly, *neither is Dr. Troia*. You tell me how likely it is that she'll be darkening Ulysses's door after that.

The glass is not two-thirds full, folks. For the faculty I've dealt with, it's often more than half-empty.

I've quoted almost all of the post because you need to think it through to see the import: That is, under normal circumstances, even "two-thirds OK" may turn out to mean that an IR manager *can't* do what a scholar actually wants—and keeping scholars interested is vital to having Green OA work. (I left out Salo's link to a SHERPA list of publishers that will support archiving PDFs, which she's grateful for. The URL is www.sherpa.ac.uk/romeo/PDFandIR.html but I've been consistently unable to reach the server.)

Peter Suber commented on Salo's post in "Talking about majority of TA publishers which permit postprint archiving" (*Open access news*, September 4, 2008). Some of his comments:

I do often say that about two-thirds of TA journals permit postprint archiving. I'll defend the claim, at least after I correct it. But first I can acknowledge that it's shorthand. When I have time and remember to add them, I add these qualifications:

First, it represents surveyed journals. Among unsurveyed journals, there are likely to be journals that do, and journals that don't, permit postprint archiving. We don't know their proportions yet. Second, the number represents journals that consent in advance to postprint archiving without requiring case-by-case requests. Many that do not consent in advance will still consent if asked individually, however. Elsevier routinely granted individual requests until mid-2004 when it decided to offer blanket permission instead. Third, it represents the journals that consent to postprint archiving, not preprint archiving. If we count the journals that consent to preprint or postprint archiving (or both), the figure rises to 93% [written in February 2006].

I base the claim on the numbers reported by SHERPA. The numbers are also summarized by EPrints. But be careful when comparing the two sources, because they use the color signifiers differently.

Dorothea is right that SHERPA doesn't survey all publishers, although it's steadily increasing the number it does survey. As of today, it surveys 418. Dorothea suggests that the SHERPA numbers are "weighted heavily" toward the "behemoth" publishers. But while there are many commercial publishers, there are only half a dozen behemoths. Any survey of 418 publishers covers far more than the behemoths. (Maybe we differ on what

counts as a behemoth.) However, it may still be true that the SHERPA numbers reflect some kinds of publishers more than others. I don't know and I'd like to know.

She says, "I have to date seen zero evidence presented that the behemoths' policies are typical of non-behemoth publishers..." If we revise this to refer to the SHERPA-surveyed publishers, rather than behemoth publishers, then I agree. We're in the same boat and we'd both like to see someone do the research. Meantime, I'm sure we both want SHERPA continue its long-term survey of publishers.

The numbers fluctuate as SHERPA adds more publishers to the survey. I used to give exact percentage figures, but when I noticed the fluctuation I decided to use looser expressions like "about two-thirds". When I started using the expression, the SHERPA percentage was between 62 and 67%. Today it's 56% (publishers allowing postprint archiving alone or both preprint and postprint archiving), which I would not say is "about two-thirds." This is a reason to downsize my loose expression (and check the latest number more often). But...

The journals published by the half-dozen behemoths may outnumber the journals published by the next 400+ publishers in line. However, the SHERPA number I've been using refers to the percentage of publishers who allow postprint archiving, not the percentage of journals. This suggests that when I want to talk about the percentage of journals, as opposed to publishers, I should revise my number upward, not downward. At the moment, I don't know how the upward correction and downward correction net out. If anyone has done the math (if 56% of *this set of publishers* permits postprint archiving, then what percentage of *the journals they represent* permits postprint archiving?), I'd like to hear from them.

One more thing I don't know and would like to know: Is the reported percentage of publishers who allow postprint archiving in decline, or is this just an artifact of the order in which SHERPA surveys publishers? Does it mean that publishers are retreating to less-green policies, or does it mean that unsurveyed publishers are less often green than the surveyed publishers?

I fully agree that the limited number of green journals, and the complex and arbitrary restrictions which even some greenish publishers put on self-archiving, are obstacles to progress.

What I notice immediately is that none of Suber's comments relate directly to Salo's argument, other than the final paragraph. Salo responded, thanking Suber for clarifying the "two-thirds" claim and getting to the real point (and here, I am paraphrasing much of what she says for reasons of length and to avoid having to provide too much backstory):

I don't think we should be highlighting [the two-thirds claim (much less the derived claim "Green OA is compatible with TA journals") *regardless of its truth value*, and my Achaea U parable was intended to demonstrate why not: publisher permission to self-archive *is not filling repositories or otherwise furthering green open access*, and that being the case, we have no reason to celebrate it, or to continue pursuing such permissions from other publishers.

The name of the game is to get the scholarly literature onto the web—that is, to make it openly accessible. Salo doesn't believe that getting a publisher to permit self-archiving achieves that aim, and believes that claiming victory for getting such permissions undermines the credibility of OA advocates. She even suggests that big publishers may be granting blanket permission as a cynical ploy, imagining a hypothetical strategist:

"Sure... let the repository-rats have their preprint and postprint permissions, why not? It's hard enough to pry a usable manuscript out of faculty when they want it published. They think some poor schmuck *librarian* can do it? Hah. We haven't conceded *one damned thing*—nothing whatever of substance will become open access because of this—but we smell like roses anyway. Win!"

Why would publishers believe this? Because faculty don't manage their own publications effectively (do you keep copies of all your publication agreements? I sure don't), practices tend to be even worse when computers are involved—and, as a result, "*most preprints and postprints are lost forever.*" Also, as creatures of habit, many faculty will happily sign over rights to publishers (they've always done it, so why not?) but will "scrutinize an institutional repository from every conceivable angle—from an initial instinct of distrust, to boot—and most will refuse to bite."

The result? Go back and read the parable. The faculty member assumes most of her publications would go into the IR easily and without fuss—and when that proves to be a false assumption, the repository librarian gets the blame. "Sure, sure, it wasn't entirely Ulysses, it was also the green-OA movement trumpeting the "two-thirds" line, and nobody meant any harm by it—but all she sees is Ulysses and his perfidy. Poor Ulysses, his own movement has hung him out to dry."

Again, there's more to the post—but that's enough of the sequence to get the overall message. Which, I believe, is this: In the real world, "green OA" is a whole lot tougher than it seems—permissions don't constitute deposits and may not do that much to facilitate them.

Will OA progress lead to Pyrrhic victory?

This time, instead of using a Salo title to lead into a Suber piece, I'm using a Suber title—a March 11, 2009 *Open access news* post—to lead into a Richard Poynder article it references. The Poynder article “Open Access: Whom would you back?” appeared March 10, 2009 on *Open and shut?*—Poynder's valuable OA blog.

Poynder's article raises a colorful question:

Should Green and Gold OA be viewed as concurrent or consecutive activities?

Poynder notes that, as you'd expect, Suber regards them as complementary (not competitive) activities, thus concurrent. And, to be sure, that “self-styled archivist Steven Harnad” regards them as competitive and argues—strenuously and ever so repeatedly—that Gold OA should only be considered *after* Green OA succeeds (which he calls “inevitable”).

Poynder reads Harnad as saying that only Green OA can resolve the affordability problem, through a twisted progression that makes sense only if you allow a huge set of assumptions (such as that reducing publisher costs will reduce publisher prices!).

Poynder suggests that, rather than Green OA eventually creating Gold OA by default, it could work the other way around: “Gold OA could marginalize, and eventually overtake, Green OA.” He then notes hybrid OA as a Trojan horse, enabling publishers to appropriate OA as their own and “defang the OA movement at large.” And notes that publishers accepting Green OA have also been insisting on embargoes—sometimes fairly long embargoes—and resisting effective OA mandates such as the NIH policy.

Poynder seems to suggest that Jan Velterop has been an effective agent for “ambushing the OA movement” through overpriced and institutionalized Gold OA, in more than one position. He concludes that this series of events (and more) can lead to publishers embracing OA while retaining (or enhancing) their profits—at the expense of libraries and other institutions. Meanwhile, green OA seems stuck at about 15% in the real world—never mind the permissions.

Poynder says “It is hard not to conclude that while the research community may win one battle (*access*) it is set to lose another (*affordability*), and in so doing lose the war.” Why? Because the same publishers will continue to control peer review at locked-in high prices. He suggests the need for a “more radical revolution” than OA.

After reading the whole article, I'm less than convinced by Poynder's reasoning, but it's clear that he's coming down on Harnad's side:

If Green OA wins the race, the research community can hope to finally free itself of both the access and affordability burdens that have for so long dogged it, and publishers will be forced to give up some of their profits. The research community will have won the war.

However, if Gold OA wins the research community will have freed itself of the access burden, but failed to free itself of the affordability burden. Publishers will have won the war.

Comments on the article itself are interesting, including a truly odd one from Harnad speaking of his many disagreements and noting the fine overall quality—as he should, since Poynder comes down on Harnad's side! Suber offers a long set of notes on Poynder's article—close to 2,000 words, unusually long for an OAN post. Excerpts (omitting some important but long arguments):

“[I]t seems that Gold OA could marginalise, and eventually overtake, Green OA.” I didn't see the argument for this conclusion...

“Hybrid OA was set to become...a tool that would enable publishers to infiltrate the movement, and appropriate Gold OA. And today it looks as though it could defang the OA movement at large.” I saw no evidence for this statement either...

“[W]hile most subscription publishers had by now agreed to sanction author self-archiving (for political reasons alone), they invariably insisted on an embargo period, from six to twelve months, sometimes longer.” I believe this is untrue. Publishers who insist on an embargo for green OA are still a small minority of publishers who allow green OA...

Suber doesn't buy terming institutional Gold OA arrangements as an “ambush [of] the OA movement” and won't accept that such arrangements are overpriced without evidence.

While we explore the many gold OA business models, and look for ways to make the revenues high enough to cover expenses without excluding authors, it's critical to remember one thing. Green OA doesn't face these problems, can be achieved faster and at lower cost than gold OA, and is not undermined by the progress of gold OA.

Richard seems to agree on the urgency, speed, and efficiency of green OA, but he seems not to agree that it's under no threat from gold OA. Indeed, like Stevan Harnad, Richard may think the virtues of green OA make it unnecessary to pursue gold OA at all, or unnecessary to pursue gold OA until green is further along. But that's where we diverge. We should pursue both at once, and I still haven't seen a good reason not to...

It's relevant to point out here that most OA journals charge no publication fees or institutional memberships at all. I've argued that even fee-based gold OA is

not the threat that Richard seems to think. But even if I'm entirely wrong about that: fee-based gold OA is a minority of gold OA, and no-fee gold OA doesn't pose any of the threats that Richard describes.

One general point in conclusion: I never saw the need to distance the access problem from the affordability problem. It's true that they are separate problems in the sense that we could solve the access problem without solving the affordability problem (e.g. with expensive OA journals). That is the prospect which alarms Richard. But the urgency of solving the affordability problem has given the OA movement some of its most stalwart allies and most enduring incentives. If Richard is saying that we should address both problems at once, I fully agree, though we may differ in some of our reasons. We should address both at once in part to avoid the Pyrrhic victory Richard describes, in part to recruit and retain indispensable allies, and above all to apply a very elegant solution (complementary green and gold OA) to a very serious problem.

Elsewhere, Ivy Anderson of the California Digital Library refutes some of Poynder's comments (not quoted here) about UC's Springer institutional deal.

Sigh...

The next piece in my lead-sheet stack is from Peter Suber, but it consists of unannotated excerpts from another Harnad screed arguing that we must not "squander" funds on Gold OA when all efforts should be spent on *universal Green OA mandates*. (There's something about "universal mandates" that makes me squirm, frankly; it's the kind of thing that doesn't sit well in an even semi-democratic society.) But I'm going to leave this one with just one quote, typical of the reason some—including repository managers—despair of Harnad's effect on OA and repositories in general:

Green OA self-archiving can be accelerated and scaled up to universality (**and this can be done at virtually zero cost**) by the research community alone... (Emphasis added.)

Sigh. As Harnad himself said years ago, repeating the same statement over and over (and over and over...) again does not automatically make it true. Dealing with Harnad, even indirectly, helps push some of us out of discussing OA altogether, from sheer frustration if not antipathy.

OA and strategy

Bill Hooker adds an interesting and important voice to the open access discussions, particularly taking on sloppy research being done elsewhere and bringing a scientist's mindset into play. His blog is *Open reading frame*; you'll find it at www.sennoma.net (it operates

under a CC 0 license, that is, "public domain as far as I can make it so"—I could legally, if not ethically, quote his entire posts without even giving him credit).

This post appeared on June 19, 2009. He cites an exchange of posts related to Stuart Shieber ("tireless architect" of Harvard's self-archiving mandate) and his speeches on open access, related to the five-institution Compact discussed earlier in this article. You can guess what Harnad has to say: "it is still a strategic mistake to focus on journal economics, and on new 'compacts' for university funding of OA journal publication fees, instead of stressing the all-important priority [that is, Universal Mandated Green OA]." There's more, to be sure (1,618 words in all), but that's the boldface, repeated message.

Shieber responded indirectly, with "The argument for gold OA support," posted June 11, 2009 at *The occasional pamphlet*. Shieber raises two questions: "Are green and gold open access independent of each other? In particular, is worry about gold OA a waste of time, and are expenditures on it a waste of money?" and explains why he believes the answer to both questions is **No**.

Shieber's essay is, while (inevitably?) briefer than Harnad's note, still fairly long (1,350 words) and rich enough in argumentation that it needs to be read in the original (direct URL: blogs.law.harvard.edu/pamphlet/2009/06/11/the-argument-for-gold-oa-support/). He explains why it's inadequate to assert that universal mandates would make affordability less urgent and even *more* inadequate to claim that "supply and demand"—that is, the market—will solve affordability problems.

Shieber deals with the real worry of scholars that, if green OA puts publishers out of business, effective peer review and branding might also vanish—and finds that one way to assuage this worry is by providing a "reasonable alternative business model": Gold OA.

In summary, a university that commits to the open access compact will more easily be able to answer objections against green OA policies specifically because it has an approach to long-range support for gold OA publishing, not in spite of it. The two models are inextricably tied. I, like Professor Harnad, am interested in facilitating the adoption of green OA policies. I proposed the open access compact in large part because I expect that adoption of the compact will lead to more green OA policies. The open access compact is therefore contributory to the promotion of green OA, not a sidetrack to it. I of course encourage universities to adopt green OA policies before gold OA support, but given that dystopian fears of faculty are preventing adoption of such policies, an open access compact that might assuage these worries should not be delayed.

He also argues against Harnad's claim that *any* support of Gold OA is "a needless waste of scarce research funds"—and that's a detailed discussion that needs to be read in the original.

The fourth of six comments comes from (who else?) Steven Harnad, beginning with an ALL-CAPS SHOUT and linking to a full response while providing a mere 400-word summary. It's classic Harnad, although it makes clearer that Harnad now *assumes* "institutional subscription collapse" as the mechanism for Green OA (eventually, maybe) improving affordability. He also calls payment for Gold OA "incoherent." The full Harnad response? A classic, more than 4,400 words (that is, more than three times the length of Shieber's essay), and full of Harnad's standard arguments. Every time Harnad says green OA is "cost-free," another IR manager loses their wings...

Which, circuitously, brings us back to Bill Hooker...and I find that my best course is to point you to the post itself. He sees strategic errors in Harnad's lengthy set of assertions and conclusions and concludes that "the optimal strategy seems to me to be the one adopted by Harvard: a Green OA mandate and careful (fiscally responsible) support for Gold OA." There is, *of course*, a lengthy comment from Harnad. And, at that, I'll let the colors fly...

Numbers

Numbers mystify many people in any case, and the numbers surrounding OA can be more mysterious than most. Phil Davis managed to make a "case" for Gold OA being a terrible deal for universities by selectively choosing numbers and universalizing them. Publishers have come up with outrageously high fees for Gold OA by equating their **revenues** with **costs** and dividing by—well, in some cases even a smaller number than the actual number of published articles, apparently. (Some Gold OA publishers seem to justify what appear to be very high fees with no clear numbers at all.)

But even the non-mystical numbers are fuzzy at best. For example:

- If it's true that 20% (25%?) of scholarly journals are now Gold OA (is that true?), is that a more or less useful number than the percentage of *refereed articles* published in Gold OA journals?
- Now ask the same question for non-fee Gold OA journals: What percentage of refereed articles (or of refereed articles published in Gold OA journals) appear in non-fee journals?
- How about the same questions for "toll access" journals?

- Has anyone done a coherent correlation analysis between impact factor and author-side fees? Between subscription prices and impact factor?

There probably are good answers to some of those questions, but as a part-time OA independent and observer, I've missed them. But some people are trying to be a bit more rigorous about some of these numbers, including the examples that follow.

What percentage of open-access journals charge publication fees?

Most of the questions posed above are second-order questions. First it's important to get clear answers to first-order questions, such as the one above, addressed in Stuart Shieber's May 29, 2009 *The Occasional Pamphlet* post.

In the popular conception, open-access journals generate revenue by charging publication fees. The popular conception turns out to be false. Various studies have explored the extent to which OA journals charge publication fees. The results have been counterintuitive to many, indicating that far fewer OA journals charge publication fees than one might have thought...

The first study of what we'll call the "publication-fee percentage," by Kaufman and Wills, showed that fewer than half of the OA journals they looked at charge publication fees. The figure for publication-fee percentage they report is about **47%**... Following on from this, Suber and Sutton provided a figure of **16.7%** for scholarly society journals charging publication fees.

Bill Hooker came up with a clever way of calculating a figure for publication fee percentage, by taking advantage of the publication fee metadata hidden in the "for authors" journal listings at the *Directory of Open Access Journals* to calculate the figure as of December 2007. Here are his totals:

Charges: 534 (18%)
No charges: 1980 (67%)
Information missing 453 (15%)
Total (excl. hybrids) 2967

Depending on the disposition of the "information missing" cases, Hooker's study indicates that **18-33%** of OA journals charge fees.

Shieber wrote a routine to automate this analysis as much as possible: "The method is effective, if inelegant." He revisited DOAJ on May 26, 2009—working against a substantially larger body of OA journals.

Here are the results computed by my software, as of May 26, 2009:

Charges: 951 (23.14%)
No charges: 2889 (70.29%)
Information missing: 270 (6.57%)

Hybrid: 1519 (26.99%)

Total: 5629

The numbers are consistent with those of Hooker's study some 16 months earlier. You'll see that the total number of full OA journals is up from 2967 to 4110, and the number with missing information has been halved from 15% to about 7%. The reduction in those with missing information seems to have gone more to those with fees than those without, so that the percentage charging fees is up some 5% and those not charging fees only up 3%. Again, depending on the "information missing" cases, the range of fee-charging journals is 23-30%. Assuming that the missing information cases are similar in distribution to those that were resolved over the last year, the figure would be about 27%. That leaves **73%** of OA journals, the overwhelming bulk, charging no fees.

Shieber does more than provide what appear to be clean, up-to-date figures. He includes the Python script used to do the analysis, so anyone else can replicate or update the study. (Ah, the benefits of having a computer scientist deeply involved in OA...)

Naturally, Harnad—ever out to undermine Gold OA (or Green OA if it's not done The Harnad Way)—responds by claiming that the vast majority of *top* Gold OA journals charge fees. The link he cites doesn't provide *evidence* for that assertion; it's just an assertion.

Author-side fee comparison: OA vs TA.

My sense is that Bill Hooker is doing more to provide rigor and challenge bad data than anyone else, as a recurring theme in *Open Reading Frame*. This essay appeared on June 18, 2009. It's a detailed post. Excerpts:

I've posted a couple of times about the misconception that all OA journals charge author-side fees, and each time I've mentioned the Kaufman-Wills study which found that 75% of the toll-access journals they examined charged author-side fees in addition to subscription charges. I thought it would be useful to compare author-side fees charged by OA and TA journals.

It's easy to work out what OA and hybrid journals charge; BMC maintains a detailed list of publisher article processing charges. Here are some examples:

PLoS journals charge in three tiers: **{\$1,300, \$2,200 and \$2,850}**.

BMC charges between \$1,105 and \$2,095 for most journals, and their standard charge is **\$1,470**

Hindawi charges between \$275 and \$850 for most of their journals, with a few titles up to **\$1,400**

Springer Open Choice, Wiley Funded Access and Elsevier's Sponsored Articles all cost **\$3,000**. (*cough*)

What is much more difficult to determine is how much the average author is paying in author-side fees at toll-access journals, because the charge for a given

article depends on number of pages and/or color figures, and in some cases also on whether supplementary information is included.

Below are a few examples; in each case for which I calculated a figure, I extracted the page and figure counts manually from a single issue. This is far too small a sample to be representative, but I'm just trying to get some kind of feel for the numbers. Further, the published figures I managed to find...are consistent with my "calculated guesses".

[Amended figures for NIH yield estimated average author-side fees of \$1,136 to \$1,515 per article.]

[Summarizing detailed analyses for these small samples, average author-side fees for various journals:]

PNAS \$1,446

Science \$1,019

Nature \$1,669

Cell \$2,031

Cell Cycle \$756

EMBO J \$2,974

Journal of Nutrition \$456

J Natural Res & Life Sci Education \$400

[and some official figures from journal publishers:]

Mol Biol Cell \$1,829

American Physiological Society (14 journals) \$1,000

J Neuroscience \$850 + color charges

Molecular Biology and Evolution \$922

Molecular Plant-Microbe Interactions \$1,275

The selection of journals is fairly random, just the first few that came to mind then whatever turned up when I was searching for things like "average page color charges." They range from prestige to niche, and **even the cheapest charge fees that amount to a significant fraction of Gold OA author-side fees.** (Emphasis added.)

It would be very interesting to extend this half-baked pilot study, but I think it would also be unavoidably labor intensive. Except for rare cases where publishers provide the numbers, there's really no way to calculate average author-side fees based on page and figure counts except by doing those counts for a representative sample of issues in each journal. (Perhaps a passing statistician could help me figure out what would constitute a representative sample...) Then you have to select which journals to investigate...

Details for each calculation are in the original post. A key point: These are *additional* charges—revenue to the publisher on top of subscription fees.

Cost to libraries: OA vs TA

I'll use that heading for three directly related Hooker articles—the one with this title, which appeared on June 18, 2009; an update and correction that relates

to both this article and the one just discussed (June 19, 2009); and “OA vs TA costs: I think I have finally got this straight” (June 21, 2009). I’ve blended excerpts from all three as appropriate.

In 2004, Philip Davis carried out a study of library costs in which he estimated the average subscription cost/article for a subset of ARL libraries and compared this with a range of estimated author-side fees for Gold OA, in order to determine whether libraries might pay more or less if all journals switched to OA. Here I’ve tried to update that study using information that wasn’t available back then.

Davis set the spreadsheet up to make it easy to update his assumptions and recalculate (kudos!), and Peter Suber (among others) pointed out that at least the following assumptions should be updated:

1. all OA journals charge author-side fees
2. the full cost of OA fees will be borne by libraries
3. TA journals charge no author-side fees

We now have five different studies (one recently confirmed, improved and updated) showing that in fact the majority of OA journals [70%] do not charge author-side fees...

We also know that research funders are increasingly willing to foot the bill for OA... A recent RCUK report showed that 45% of authors publishing in fee-based OA journals had their costs covered by their research funders...

I’ve updated two further aspects of Davis’ spreadsheet. First, we now have better information about the actual range of author-side fees charged by those OA journals that do charge them. Rather than Davis’ \$2,500-\$5,000 range, I’ve used \$1,300 (PLoS ONE) to \$3,000 (most of the high-profile hybrid programs). If the adjusted TA cost/article falls within this range, the prediction is that the OA and TA models cost about the same from a library point of view.

Second, Davis assumed that the scholarly literature made up 50% of library serials expenditures. I don’t know where this figure came from (the spreadsheet refers to a report which does not give any further information), but I think the real value is closer to 90%... [Detailed reasoning omitted here.]

[Summarizing, after all corrections included in the update have been applied: If libraries pay 40% of Gold OA fees, *no* ARL library will pay more for universal Gold OA than for current subscriptions—and at 20%, all ARL libraries would pay *less*. As Hooker notes, one set of reasonable calculations suggest that the actual cost could be around 13.5%]

[Under worst-case scenarios, which institutions might pay the same or a little more? Essentially, research-intensive schools that publish a lot—e.g., UC Davis, UCLA, UCSD, Cornell and a few others.]

The final article in this cluster looks at total *revenue* per toll-access article, combining subscriptions and author-side charges. Without repeating the details, Hooker concludes that it’s in the range of \$2,106 to \$2,886 and notes: “Sure puts one-time, up-front Gold OA fees in a different perspective, doesn’t it?” To say nothing of the 70% of Gold OA journals that don’t charge such fees...

Scandal!

The ever-vigilant Phil Davis, who of recent years seems to be looking for every opportunity to discredit OA, contrived a scandal and publicized it in “Open Access publisher accepts nonsense manuscript for dollars,” posted June 10, 2009 in *the scholarly kitchen*. The short version:

- As with others, Davis had received numerous invitations to publish in Bentham journals—Bentham Open, that is, a huge launch of “more than 200 peer-reviewed open access journals” from Bentham Science (which also publishes several dozen toll-access biomed and medical journals). (The count’s now up to “over 250.”)
- Davis used a program that generates grammatically correct but (presumably) meaningless articles in computer science. (The site for SCIGen, pdos.csail.mit.edu/scigen/, suggests using it to prepare submissions for “conferences you suspect might have very low submission standards”—and a SCIGen-generated paper was accepted for a conference. I’ll submit that using the program to generate an article for a journal you suspect might have very low refereeing standards is in the spirit of the program.) The program even includes figures, tables and references in the nonsense output.
- He submitted the manuscript to *The Open Information Science Journal* (a Bentham Open journal), using two generic names as coauthors and an affiliation that should have been a giveaway: The Center for Research in Applied Phrenology. (Apart from the acronym, you wouldn’t expect lots of quality computer science from an institute that studies reading personality traits from lumps on people’s heads.)
- Bentham confirmed receipt of the manuscript almost immediately and, some four months later, *accepted* the manuscript after asserted peer review, asking for the \$800 publication fee. At which point the co-authors retracted the article, citing “several errors in the manuscript.”

- Davis sees no evidence that peer review actually took place and says that anyone proficient in English (even if ignorant of computer science) would spot it as nonsense. “Had it gone through peer review, I should have received reviewer comments.”

The final two paragraphs of Davis’ post—and note that he’s tried this stunt with Bentham before, with different results:

From this one case, we cannot conclude that Bentham Science journals practice no peer review, only that it is inconsistently applied. Earlier this year, I reported on a case in which a nonsensical article submitted to another Bentham Science journal was rejected after going through peer review.

While one should be careful not to generalize these results to other Open Access journals using similar business models, it does raise the question of whether, at least in some cases, the producer-pays-to-publish model may unduly influence editorial decision-making. One may also question whether publishers like Bentham see a lucrative opportunity from the OA movement, considering that academic libraries are establishing author publication funds to pay Open Access charges. [Emphasis added.]

The payoff here is the first sentence in the second paragraph—and, given that 70% of OA journals *don’t* require author-side fees while 75% of toll-access journals *do* require author-side fees, my immediate rejoinder would be “Since most toll-access journals require author-side fees *and* since toll-access publishers justify increased subscription prices by pointing to increased publication volume, one must also raise the question of whether such fees may unduly influence editorial decision-making in traditional journals.” (David Prosser offers a more elegant version of that wording.)

Are there toll-access journals that will publish pretty much anything they can get their hands on, with or without author-side fees? Almost certainly. (In one comment, Davis notes the “long tail” of subscription journals “that exist because they fulfill the need for second-tier faculty to publish their articles and to sit on editorial boards”—even though the papers will rarely be either cited or even read.)

Has this stunt ever worked with traditional journals? It has, although admittedly within scholarly fields where grammatically correct nonsense may sometimes be hard to separate from scholarship. (Apparently, it’s also worked in computer science—in an Elsevier journal, no less, with a paper created by SCIdgen.)

The long set of comments on Davis’ post includes those who had bad feelings about Bentham *as one pub-*

lishing company—and others who saw this as reason that “the pay to publish model has to end.” (Does that also go for author-side fees in subscription journals? If so, wave goodbye to most journals.)

When David Prosser raises a rejoinder similar to the one I raised, Davis responds that *librarians* prevent crappy subscription journals from succeeding. That may be true if the crappy subscription journals are high-priced, but I’m certainly aware of publishers who’ve made a business out of salami-slicing mediocre journals that only cost a few hundred dollars each, and tend to fly under the radar at larger institutions.

Gunther Eysenbach notes that he’s warned people about Bentham before, that there’s now an Open Access Scholarly Publishing Association (which Bentham has not joined) to maintain quality standards in OA publishing, and that it would be interesting to have a control test—that is, to send fake papers “to small subscription-based journals which are desperate for submissions.” Davis mentions existing methods publishers use to cope with a lack of good papers—but repeats that librarians will spot this and cancel the journals. I dunno; I don’t see that happening in moderately-priced journals. (Librarians in the group can probably name a few “quarterly” journals that rarely produce four issues a year and still seem to have plenty of subscriptions.)

A rather lovely comment from Elizabeth Davies notes that “some peer reviewers might not be capable of recognizing gibberish, because that is what they themselves produce.”

As you might expect, Peter Suber cited the post and offered informed comments. There’s been suspicion about Bentham for more than a year. The hoax certainly demonstrates incompetence at Bentham—but the question is whether something larger is at work. While Davis cautions against generalizing, he does so in a way that suggests generalizing, sort of the “don’t look over here, there’s nothing to see” approach.

The thing is, Suber and others interested in Gold OA are the biggest enemies of scam artists within the field. Dishonest and incompetent OA publishers weaken OA, possibly more than dishonest and incompetent toll publishers weaken traditional publishing.

The followup, across many different blogs and other platforms, was interesting. Bentham itself claimed that it *knew* the article was a hoax and was only pretending to accept it in order to learn the authors’ true identities. That, frankly, doesn’t fly; Davis notes that nobody attempted to contact him directly.

The editor of the journal in question, *who had never seen the article*, resigned. So has at least one edi-

torial board member of another Bentham Open journal, noting that in his time on the board he had *never* received a paper for review. Another Bentham journal editor resigned when a paper (asserting that the Twin Towers attack was a conspiracy) was published without her knowledge or approval. (The editor also said that the Davis/Anderson hoaxers were guilty of somewhat unethical behavior, since the review process is based on trust.)

Summing up:

- There's at least one questionable OA publisher (and probably others). For that matter, *any* time a publisher announces scores of new journals all at once, whether toll or free, I'd be suspicious.
- So far, there are at least two known examples of nonsense papers being **accepted** by subscription journals and two known examples of nonsense papers being submitted to a known-to-be-questionable OA publisher—and in one of those two cases, the paper was rejected. How's this for a conclusion: "Based on a small-scale sample, Gold OA publishers are only half as likely as subscription publishers to accept hoax papers." But, you know, the plural of anecdote is *still* not data—in fact, these four instances say *nothing whatsoever* about relative standards of peer reviewing across the journal literature.

Ah, but there may be other lessons to learn.

Opportunity in opprobrium

That's Dorothea Salo's suggestion in a June 13, 2009 *Caveat Lector* post. Excerpts:

[O]ne way to look at this unpleasant situation is as an information problem. If that suggests to you that I think librarians have a role in solving it, you know me entirely too well. In fact, I think we have to get a handle on it, because we are and will continue to be some of the organizations funding gold OA. Imagine the mess, if a well-regarded academic library funneled money to a Bentham!...

...But taking a stand gets sticky, too, because...the last thing any academic service center wants to get involved in is telling faculty where they can and can't publish. As gold OA takes on increasing importance, anyone with funds to disburse toward author fees may well land—or be perceived as having landed—in precisely that position. How do we even begin to think about that?

Well, one way is to think of ourselves as research funders, not unlike the NIH or the Wellcome Trust. If we're paying the money, we deserve a say in where it goes, and we're well within our rights to say that [sketchy publishers are] right out. As librarians, we make collection-development and purchasing deci-

sions based on assessment of information quality, right? (Yes, yes, "when not prevented from doing so by Big Deals and similar less-than-savory practices," granted.) This is the same thing, just at a different point in the process. It shouldn't be a problem.

Of course, I've just begged a huge question. How do we know [who's sketchy]?...

I think OASPA's response to the Bentham situation points to part of the way forward. If OASPA membership becomes a seal of approval for all-OA publishing operations, then it's dead simple for any library that funds author fees to hold to a policy of "if it ain't OASPA, we ain't paying." This puts a significant burden on OASPA, I grant you—if nothing else, they have to have the guts to kick out a bad apple—but my sense from that post is that they're at least willing to consider picking up this gauntlet. If so, good for them.

I'm not sure that OASPA membership solves the entire problem, unfortunately... and I don't know anything about OASPA's membership structure or finances, so I apologize in advance if this line of thought is completely misguided... [See the post itself for an expansion of this discussion.]

Even with...caveats, I think an OASPA certification program represents a tremendous opportunity for the OA community. Gold OA is still small. It's much easier to put meaningful quality regulation in place over a small, emerging, prestige-hungry industry. If gold OA manages to do that, then it suddenly has another competitive advantage over toll-access, which hasn't done so and (given its extent and decentralization) very likely can't...

Go read the whole post; there's more there worth thinking about. You might also read "How to avoid questionable OA publications?" in *Be openly accessible or be obscure* (tillje.wordpress.com/), but Jim Till's suggestion really only works for the health sciences.

Framing and Mysteries

"Framing" is an interesting word. Some might say Phil Davis has been trying to frame the OA movement (in the "I been framed!" sense) for some time—publishing cost analyses using a wildly biased set of assumptions, engaging in hoaxes aimed at exposing shoddy OA peer review (and "non-suggesting" that there might be a general problem), and so on. And he suggests mysteries where there may or may not be any—which we'll get to in a moment.

He also discusses another kind of framing, which I'll call rhetorical framing, in "Framing the Open Access debate," posted March 3, 2009 at *the scholarly kitchen*. Right there in the title is an important piece of framing: asserting that there is a legitimate debate.

Is there? Is there a legitimate argument that OA *in general* is a bad thing and to be avoided? Remember that to argue against OA you must argue against both colors—which also means you need to argue that most of the biggest for-profit publishers are *wrong* in accepting green OA. An editorial in *Communications of the ACM* questions the need for Gold OA—but ACM is a Green OA publisher (and one with generally reasonable prices). Too bad that the editorial notes the use of surplus revenues from publications to cover other activities—because some librarians (Peter Murray, in this case) are questioning this response:

I am a librarian. I view this distribution of funds as a problem. A library's primary mission is to gather, organize, and disseminate scholarly information from a variety of fields. To the extent that library subscription dollars exceed the actual cost of producing journal content to "support the rest of ACM activities" goes beyond the mission of the library.

There's a reply that begs the question of *hidden* subsidies, which these are.

Daniel Lemire asks "Is Open Access publishing the solution? Really?" in a June 17, 2009 post on his blog (www.daniel-lemire.com/blog/). The post trivializes OA's virtue ("Open Access research papers might have marginally more impact") and stresses the "costs"—which he summarizes thus:

There are far fewer Open Access journals to choose from.

On average, Open Access journals have lower standing.

There are more than 5,000 OA journals, including 155 (according to DOAJ) in Lemire's field, so one has to wonder about the first point. The second is certainly open to debate—but in that case, Lemire's only arguing against Gold OA. And, in fact, he's arguing for "disruption"—he doesn't regard OA journals as sufficiently disruptive.

What's disruptive? Well, he publishes his preprints to arXiv. Before that, he put them up on his website. This *would* be Green OA, if he was also submitting the articles to peer review, but Lemire's *really* disruptive:

Self-publishing is both simpler and more convenient than traditional publishing. **It is disruptive.** As is often the case with disruptive solutions, it lacks some important features. For example, reputation, peer-review, quality control, review, validation, authentication are difficult with self-publishing. But that is to be expected. The solution is not to try to emulate these features one by one. Indeed, we may find that many of these important missing features are not relevant.

From Lemire's perspective, maybe there is an OA debate—but hardly the one Davis wishes to frame. Davis

is constructing his own frame: Asserting that, in crafting a straightforward argument for OA, adherents have "alienated one community that should be on their side: the non-profit scientific societies" by creating a caricature of publisher behavior. Here's a paragraph that will get me in trouble:

By doing so, open access advocates ignore the fact that many society publishers are non-profit and use the surpluses from their journal to provide grants and opportunities for young scientists, fund public awareness of science programs, among other social benefits. Open access advocates ignore the fact that scientists share a common ethos of openness and a culture of sharing, and that these values are at the heart of most scientific societies.

To which I can only respond as I have in the past (perhaps too often, and this might just be the last time): *It is unreasonable to expect libraries to fund those other activities.* If societies wish to make the case that universities should subsidize association activities, that case should be made *to the departments in question.*

It is *certainly* true that many (not all) societies charge reasonable prices for their publications and offer reasonable author arrangements. It is *also* true that many societies have given over their journals to megapublishers and turned the journals into very expensive items. If the scientific societies are open and committed to sharing, they should openly demonstrate just what their actual journal costs are—and should be *leaders* in making sure articles can be shared.

In fact, however, Davis is attempting to frame a "debate" by creating his own, overstated, black-and-white generalization and opposition:

Open access advocates will continue to accuse publishers (as a group) of being uncaring and working against the public good. In turn, publishers will continue to accuse open access advocates of being irrational ideologues.

One thing is clear—this debate was never about science.

I'm trying to remember the number of times I've heard Peter Suber or Charles W. Bailey, Jr. or Dorothea Salo or Stevan Harnad accuse *all* publishers of being uncaring and working against the public good. (Yes, there are doubtless OA advocates who make such wild charges—but, you know, I've never encountered them.) So far, I come up with a goose egg—a big Zero. On the other hand, it's certainly true that *a few* publishing spokespeople—not many—speak of OA advocates in a way that could suggest "irrational ideology."

A more formal and sophisticated journal-article equivalent to this post appears in the February 2009 *Journal of Electronic Publishing* (12:1), an issue devoted

to open access. I'm not going to comment on that article or, in general, the issue, although I suggest checking the contents list and seeing whether you find some of the articles worth reading.

One commenter notes that some (not many) societies have become "major quasi-commercial publishers that happen to have memberships attached" while retaining the "fig leaf of non-profit status." Those surpluses may yield scholarships; they may also yield very high salaries. (A certain chemical group comes to mind.) Others wave the DC Principles flag (I've commented on that before)—and one raises the point I've been raising.

Dark secrets?

The title on Phil Davis' May 13, 2009 *scholarly kitchen* post is "Dark secrets: Open Access and author processing charges." To read the lead, it's a Very Big Deal—one that gets a lot smaller in the rest of the telling:

You would have thought I was requesting a field manual for interrogating prisoners of war or a list of members on Dick Cheney's Energy Taskforce. At least in those instances, I would have received a response that answering my questions violated national security or "executive privilege."

All I did was ask five librarians at institutions administering Open Access publication charges two simple questions:

"Can you provide a list of Open Access articles that you have supported through your author support program," and "Have you rejected any requests to date?"

The librarians serve in public institutions. Davis makes a point about "publication" implying "public" (so if a publication's hidden behind toll walls, it's not really a publication?) and then the "results" of his inquiry. To wit, **two weeks** after the queries, two of five responded, but without lists and actual numbers (I don't see a request for numbers in his question)—and neither said any bona fide requests had been rejected.

And Davis produces a post with a conspiratorial headline and first paragraph, then goes on about the need for transparency and "stonewalling." His close:

Those who campaign for Open Access need to be held accountable just like everyone else, and budget transparency is the first step.

Before I even look at the comments:

- *Two weeks?* Failure to respond fully within two weeks isn't conspiracy. In an institution as complex as any of those libraries, it's a *miracle* that two of the five were able to compose a response and get it approved within two weeks.

- I don't believe the libraries are the sole granting authorities in all of these cases, and I don't believe librarians could rapidly release such lists without consulting other parts of the universities, probably including counsel.
- Davis isn't calling for OA to be accountable "just like everyone else," he's calling for a much *higher* standard of accountability. Show me "budget transparency" at for-profit publishers. For that matter, show me scholarly societies who will (*or can*) break out revenues from journal publishing, *direct* costs for journal publishing, and *exactly* where the surplus goes. (Do I believe there are thousands of scholarly society honchos earning millions of dollars? Not really. Do I believe there are *some* society people with fairly lavish lifestyles subsidized by subscription fees? Well, in the absence of budget transparency, yes, I do.)

Mostly, I don't think there's a scandal here. What of the commenters? Several, *appropriately*, say this is an important conversation. "Amanda R" finds it easy to imagine that such a request would be "bumped around internally for a few weeks" and isn't quite ready to start calling this "dark secrets." Davis doesn't think it's constructive to speculate on why three libraries didn't respond rapidly—but apparently feels provocative headlines *are* constructive. Robert Kiley notes that the list of papers supported by Wellcome is readily available.

JQ Johnson points out reasons it might *not* be appropriate to provide such lists—and notes that, when he queried the same institutions, he got *lots* of cooperation. He wonders whether Davis asked the right people. Davis says this all "skirts around" the main issue and suddenly brings in another issue: "Transparency does not mean selective dissemination based on who you consider to be your friends or enemies." Since Davis does not tell us exactly who received his requests or how he determined that they were the right people to ask, this is a red herring: We have no basis for assuming Davis was denied information because he may be increasingly viewed as a dedicated enemy of OA.

Overall? There's not much of a story here, other than that big, bureaucratic institutions don't respond to requests for information instantly. The big story, if any, is the sinister tone of Davis' title and lead paragraph.

Peter Suber quotes much of the post and adds his own comments:

Phil's title and opening sentence are a little melodramatic in light of the results, but he asked two good

questions and I'd like to know the answers myself. Or at least I'd like to know how many requests the funds received (for OA journal publication fees rather than something else) and how many they rejected. I don't need to know which authors or articles were subsidized. I'd also like to know the range of fees requested, and how often the fees would pay for libre OA rather than gratis OA. I'd like to know how often the fees would go to hybrid journals with a double-charge business model (i.e. not promising to reduce subscription prices in proportion to author uptake). I'd like to know the fields or departments of the requesting faculty. If the funds have a good reason not to share anonymized data about their use, I'd like to know what it is. (This is not a hostile question; there may be very good reasons which haven't occurred to me.) I should add that I also have lots of questions about the business information at OA and TA journals.

One more question: When a survey has a low response rate, why assume that the surveyed people or institutions are suppressing information?

I don't have much to say about Suber's response. "A little melodramatic" is nicely understated; the final question is, I think, an excellent one. And, after Suber points it out, I'd agree that a *list* is probably not an appropriate response—and, as a question, almost seems designed to make responses slow and difficult.

Bill Hooker (in a May 26, 2009 *Open Reading Frame* post) also notes that Davis "plays the conspiracy theory card way too hard for my taste" and that, when he left a comment asking pointed questions about Davis' methodology (e.g., who did he contact, did he follow up, how does he know the email wasn't trapped as spam)—well, after almost two weeks, the comment not only hasn't been answered, it didn't even appear on the blog (then or now, four months later):

What dark secrets is Philip Davis hiding? What dim, Crotty-esque ambitions of being the famous naysayer, the Nicholas Carr of Open Access, are forming even now in the troubled subconscious of this —

Or, you know, I just got stuck in the spam queue. It happens. :-)

He also notes that the "everyone else" Davis claims need to be held accountable should include "bloggers who wish to hold librarian feet to the accountability fire." Davis certainly didn't make his research methodology transparent, and that's important. Appropriately, Hooker mentions Hanlon's Razor: "Never attribute to malice that which can be adequately explained by stupidity." Substitute "bureaucracy" for "stupidity" and this seems to fit perfectly. (I've always preferred "incompetence" as the final word in Hanlon's Razor, as lots of incompetent people aren't stupid at all.)

A FriendFeed discussion includes some useful notes. Christina Pikas says "the worst part of this is figuring out who you would send a request like that to" and "Almost seems that he's taking confusion for malicious intent." Indeed.

The Problem(s) with Green OA

Maybe the heading for this section should be "IRs"—and maybe I should be sensible and drop this section entirely. What do you do when one path toward better access relies heavily on a specific instrumentality—and the loudest mouth advocating that path consistently, um, fails to appropriately represent (the kind wording) the reality of that instrumentality? Maybe, if you're an outsider like me, you *read* the stuff Dorothea Salo's written in the past, add a few other things from John Mark Ockerbloom, Alma Swan, Richard Akerman and others and choose not to comment on it.

Or maybe you let some particularly juicy or interesting pieces accumulate—mostly from *Caveat Lector*, a few from elsewhere—until you have (good grief!) *eighteen of them*, exactly two-thirds from *Caveat Lector*, and attempt to summarize them coherently. Which is what this particular section is on about, although coherence may be lacking. I'm *not* including any posts from Green OA's chief advocate. For one thing, all of them seem to be the same message with slightly different wordings; for another, he does enough harm to his own movement all by himself, without my assistance. Indeed, the very first piece I'm discussing here kicks off with an entirely appropriate statement:

With allies like Dr. Stevan Harnad, do institutional repositories really need enemies?

That's the start of "Allies," posted July 9, 2008. There's a link to a Harnad piece, beating up on Nature Publishing Group (NPG) for offering to take care of depositing postprints into PubMed Central. As is typical with Harnad, he equates self-archiving to "a few keystrokes" and harangues any OA solution that is not, precisely, 100% what he advocates.

It turns out that NPG is willing to deposit postprints in institutional repositories (henceforth IRs) as well, which negates some of Harnad's criticisms—but Harnad still insists that "If Nature really wants to help OA, then dropping its access embargo would be a lot more helpful than saving authors from having to do a few keystrokes."

Some of what Salo says about that and about NPG's offer:

Wait a second. I thought keystrokes were the big limiting factor in reaching green-OA's holy grail. A lot

of keystrokes were just eliminated, and very likely more will be as other publishers (who watch NPG like hawks, because NPG is amazing) follow suit. Surely this is a good thing? Guess not.

Truthfully, though, it's not the count-your-blessings aspect of this silly little kerfuffle that gets under my skin. It's the pattern, the pattern of OA advocates thoughtlessly backstabbing their allies, with IRs being a favorite whipping-boy, and it's of *long* standing. And yes, before you ask, I'm as guilty as anyone when my annoyance with DSpace boils over.

Dr. Harnad pillorying NPG for helping IRs stabs me and my fellow repo-rats in the back, because we *need* cooperative publishers and impetus toward interoperability. Noisy proclamations about how everybody wants open access likewise backstab repo-rats, because if library administrators believe this bushwa, they inevitably blame their local repo-rat for still-patetic adoption and content-capture rates. Tarring all librarians, repo-rats included, with broadstroke brushes about nitpickiness and get-with-the-program-already—well, I need say no more. Ignoring librarian contributions to OA, particularly green OA, is worse. All of this harms IRs. If IRs are still important to OA (which is, I grant, arguable), it harms OA.

So could we all stop it, please? Myself included. We're allies; let's act like it. At the least, we could start considering the impact of our careless words on our allies.

Can't we all just get along? A great idea—but one that's unlikely to reach fruition, I'm afraid.

A few weeks later, after Salo gave a keynote at a conference on repositories, she posted "Repository tidbits" (August 5, 2008). A few pertinent excerpts:

The Edinburgh experience left me with a complex observation that I tried unsuccessfully to get across to one person who was pressing me *really hard* to identify "successful" repositories in the United States. (No, it wasn't "define success," though that was a tempting question. The reality is that success for institutional repositories is defined—by the people who *matter*, anyway—in terms of items deposited.) It's an important enough point, and I made it inarticulately enough at the time, that I think it deserves a piece of a blog post.

...There's an understandable tendency to look to "successful" repositories for examples the rest of us should follow. Nothing wrong with that in and of itself. Problems arise, however, when one assumes first, that unsuccessful repositories have nothing to teach; second, that the successful repository experience is the norm; and third, that the locus of success lies within the repository and its staff.

Because it doesn't...

Successful repositories have sufficient backing from their libraries and their university administrations to make some-

thing work. I can't make it any simpler than that. Without that support, the best repository-rat in the world will not succeed. With it, you don't need an Einstein.

Exactly *what* successful repositories make work varies quite a bit, according to the talents and creativity of the staff involved and the nature of the support provided. This is why it's impossible to write the "winning recipe" for a successful IR...

The other thing that successful repositories have is *leave and resources to experiment.* They have to. The standard repository software package, as I have argued ad nauseam, is wholly inadequate to fuel a successful repository program. This means that the well-dressed repository manager has some combination of elbow room in her job description, developer time, student help, librarian alliances, and administrative weight to work with. Again, the exact combination will differ from institution to institution—but a manager without any of this might want to rewrite her résumé before her current job tarnishes it.

So much for the successful repository. Let's talk for a moment about the *typical* repository. If either of my two(-plus, if you count my contribution to a consortial IR) repo-rat positions is any example, the typical repository is running on a wing and a prayer and the dedicated efforts of *one FTE or less.* I say "dedicated" for a reason, because it's next-door to impossible to garner impressive results from the voluntary efforts of scattered librarians who don't have any kind of imprimatur from above for their repository efforts.

This is no way to run a program, folks—and yet as best I can tell...it's the *typical* way it's done. These repositories cannot be successful...yet they are the majority case! What does that say for the repository movement?

The thing is, just telling repository managers "You need developer time, student help, willing librarians, and administrative support" is a *useless way to behave.* This message doesn't need to go out to repository managers; *we know already,* and there's nothing we can do about it. Again, we don't have the *administrative support* to garner all that other stuff. The message needs to go to *research-library administrators,* many of whom have yet to hear it. We have sufficient reasonably successful examples now that we can say this with authority. So let's.

What may *not* help this effort: A high-profile person who repeatedly insists that IRs *don't really cost much of anything*—certainly not the salary of one FTE!

A month later, Salo posted two successive posts, "Feeding Mr. Blue" on September 9, 2008 and "What do we want from IRs, and what are we doing to repository rats?" on September 10, 2008. I'm picking out small excerpts from both. (Mr. Blue? An invisible hero who figures into several of Salo's IR discussions.)

The key statement in the September 9 post: **Voluntary unmediated self-archiving is not a viable model for institutional-repository population.** Which is *not*, as Salo goes on to explain, the same as saying “IRs are not viable.” So what’s the difference?

- **Voluntary:** You need a mandate—but probably *only the faculty* can establish an effective mandate (and you probably need a library mandate first!). If a mandate’s not possible, bribery might work.
- **Unmediated self-archiving:** Yes, faculty save their papers—but they save stuff all over the place. They’re not likely to put it in the IR on their own, but mediated deposit might work. (Salo has a fair amount of detail here, much more than I’ll quote.)

That summary certainly doesn’t do justice to the useful post. The second post is longer and even more difficult to summarize easily, but it boils down to suggesting questions an institution needs to answer before beginning an IR.

There are still, apparently, institutions who believe in the “build it and they will deposit” model. To Salo, the key question is this:

What content do I want from this initiative, and what am I willing to do to get it?

And if the answer to the second part is “I’m willing to run and market an IR,” Salo says you shouldn’t start one—it’s not enough and you’ll waste people’s time and money. Instead, she suggests, the focus should be *not* on IRs as such but on the content types desired and how you’re going to get them.

Once we focus on the *stuff we want* instead of the place we’re going to put it, we open up the questions we should have been asking all along. How does this stuff get produced, and how could we help produce it in a way that keeps it available to us? What happens to it when it’s done? What incentives can we offer to have it given to us, and are those sufficient to counter any opposing incentives combined with natural inertia and the actual difficulty of the task? Failing that, how do we find out about the existence of the stuff we want, and how can we then get our hands on it in the form in which we need it?

Given all the effort needed to do what an institution wants, is it still worthwhile? Salo says “**No**” may be the right answer in some cases. When failure is predictable, maybe it’s not worth the attempt. There’s more, specifically about the damage done to professionals by saddling them with unworkable IR duties, but this relates back to this section’s overall topic at the end of the post:

Until the open-access movement turns honest about the labor required to accomplish its real goals—notably, we’re fairly honest about gold OA and mealy-mouthed still about green—and acknowledges the damage it has done to the labor it’s already mustered, it can’t make a better start.

Let’s move away from *Caveat Lector* for a bit. John Mark Ockerbloom posted “Repositories: Benefits, costs, contingencies (with an example)” on September 11, 2008 at *Everybody’s Libraries*. He begins from the same perspective as Salo—“Without a clear sense of benefits and costs, you won’t have a sensible repository strategy... You have to go in knowing what you want, and being realistic about what you’re willing to invest to produce it.” More:

Even when your initial plan is sound, you have to be prepared for change, and the unexpected. Technology changes quickly. Online tools, communities, and scholarly societies also change. Methods of scholarship also change, often more slowly, but sometimes in significant ways. Even if you’ve done your homework, you may eventually find that the repository that seemed just fine a few years ago doesn’t really meet your needs like it used to. Maybe the software hasn’t been updated as you’d like it, and there’s a better system available now. Maybe you’re storing different kinds of things, or you’ve found a new application that your scholars really want to use that’s not compatible with your existing setup. Maybe the formats you’re managing have gone out of date. Maybe it becomes more cost effective to move to a big externally managed repository that your scholars are flocking to already— or away from one that they’re not finding useful. Maybe you even decide it no longer makes sense for you to maintain a particular repository.

You need to start thinking about strategies for change (and for exit) the moment you start planning a repository. Remember, repositories ultimately don’t exist for themselves, but for their content (and for the people using that content). And the kind of content that libraries often care about is likely to remain relevant much longer than any particular repository configuration. You want to ensure that the content remains useable for as long as your patrons care about it, even as it moves and migrates between systems (and possibly, between caretakers).

Ockerbloom’s example, at Penn, is planning for *data* repository services. He details some of the questions and costs they need to consider.

Contingencies, by their nature, tend not to be fully foreseeable. But there are a few obvious things we can ask about and plan for. Will our data still be readable for decades to come? Can we migrate it to new formats, and if so, what would be involved? Can we make sure we have good enough metadata and annotation to

know how to read, use, and migrate the data in the future? Do we have clear identifiers for our content that will survive a move to a new platform (and leave a workable forwarding address, if necessary)? What happens to our content if our repository loses funding, our machine room is sucked into a mini-black-hole, or we simply decide it's not worth the trouble of keeping the repository going? What do we do if we're told to withdraw or change the data we're maintaining, by the person who deposited it, by someone else using or mentioned in the data, or by the government? We won't necessarily come up with definitive answers to all these questions, but brainstorming and thinking through possible and likely scenarios should help us know what to expect and reduce the chance of our getting caught unawares by a costly problem.

As he points out, if you just put up an IR without thinking about those issues, faculty are less likely to think that the IR's benefits are worth their time.

Running a large, successful, long-lasting repository takes a lot of work over its lifespan. Better to do some planning work up front than get stuck with a lot of costly and unnecessary work later on.

Next we move to *Confessions of a science librarian* in its old home, jdupuis.blogspot.com, for an October 5, 2008 post—but we get no further away from Dorothea Salo, since Dupuis' post is an interview with her. A few choice excerpts from a 3,200-word interview:

I think the institutional repository was a noble and worthwhile experiment, but as a tool for redressing the imbalances in the scholarly-communication system, it is a failure...

I think we libraries have a lot of market power that we are not using properly. I've heard publishers talk about their industry, and what they invariably say is "we will follow the money." That means libraries; as individual subscriptions dwindle, WE are the ones with the money. They'll follow us—but we aren't leading them toward open access. We're squealing like stuck pigs about the stalemate, yes, but we're not reallocating any of our serials funds to support gold open access. I think this is a serious mistake...

...I want to see us cancelling overpriced journals, regardless of their impact factors or usage statistics, and standing up to faculty when they ask why. We need to say "no" loudly and clearly more often, and we need to divert some of the serials money we save thereby to gold open access. (Some should go back to monographs, of course.)

In response to the question "If you could get one message across to faculty at your institution about Institutional Repositories, what would it be?"

Faculty specifically? "Let me help you."

Frankly, though, a more productive message would be directed at my fellow librarians, and would read "I can't do this without you. Help me."

On the future of IRs over the next five years, looking for best and worst case scenarios:

Worst case is easy: they are defunded and die... If the software remains obtuse and difficult, if the goals remain socio-culturally impractical, if the services remain under-resourced and poorly understood, IRs are doomed. At a good many institutions, I believe this is inevitable...

Best case: IRs shift from "warehouse at the end of the digital train tracks" to a set of services and systems that manage, safeguard, and shepherd the digital products of the research process all the way through, soup to nuts...

On the future of journals, publishers' business models, open access and all...

...The publishing lobby will continue its stunning mendacity, largely though not entirely unopposed by rank-and-file publishers. There will be more open-access journals. It is likely to become harder to assert that open-access journals are unsustainable, but that won't stop the publishing lobby from trying—and it won't stop a few gold journals from folding, either. We will continue to argue about citation advantages, and just what a citation is worth. Faculty will continue to feel whipsawed by all this.

Which will bring us back to Do...or, rather, to *Caveat lector*. "Ya think?" on October 24, 2008 notes a RAND Corporation report on UK repositories that says things Salo's been saying, for example:

[D]igital repositories are currently underutilised, and...there are significant barriers to a strategic commitment...

Even if most of the barriers identified in this report—e.g., the lack of awareness, a technology that is in its infancy, risks of reputation damage, or the administrative burden of depositing—can be overcome, one major challenge remains for digital repositories, namely the lack of incentives for the wider institutional community to provide content for these repositories...

...There appears a misalignment between the objectives of the repository and the needs of different groups of stakeholders.

After these quotes and a few comments, Salo gives more specific advice to most academic institutions than I'd seen before:

If you aren't a doctoral institution, don't bother with an IR. No, I don't care what Harvard did; you aren't Harvard, and you have many less-futile things to do with your precious library budget and staff.

Speaking of smaller institutions and IRs, Salo asked about IRs at small institutions—and got some res-

ponses, all requiring anonymity. Some conclusions (from a November 3, 2008 post):

Some of these institutions are using IR software as a quick-and-dirty digital preservation mechanism, mostly for institutional records. These institutions don't have any particular commitment to open access, nor are they under any illusions about their faculty's commitment thereto. They'll let faculty deposit if they've a mind, but that's not what the IR is there for. A variant on this scenario is IR software considered alongside other content- or knowledge-management systems.

So far, so good; this seems decently sensible to me, though in all honesty I can't imagine most IR software standing up well against a real CMS. I'm relieved, though; I appreciate knowing that my profession mostly has its head on straight.

Unfortunately, that's not the whole story. At least some of these small institutions have The Shiny in their sights. *Everybody* has an IR these days, so why don't we? If we build it, they will come!

In at least one such place, I am told, cooler heads higher on the totem pole are prevailing, which is all to the good. In at least one such place, the hotheads are the top of the totem pole, sadly.

To be fair, small institutions have one major advantage large ones don't: scale. Yes, scale. Small is beautiful; at many small schools, it's perfectly feasible for the library to have one-on-one relationships with nearly every researcher on campus, especially given that teaching-focused schools aren't necessarily publishing all that much research! One-on-one outreach works; it's nearly the only technique that *does*.

That said, I see a Sisyphean climb even for small schools. Most importantly, small schools don't have many librarians, and those librarians don't have much time to spare. Forget about digitization or mediated-deposit programs; they're wasteful luxuries at a small school. Many such libraries, even many small institutions, don't have nearly the technical savvy it takes to run even EPrints, which is the most straightforward of the Big Three open-source IR packages...

There tend to be cultural differences between small teaching-focused and large research-focused institutions as well, ones that do not bode well for open access. One is obvious: these places tend not to have physicists, medical researchers, computer scientists—precisely the disciplines that are early adopters of open access. What they have in plenty are humanists and social scientists, who see OA (when they see it at all, which most frankly don't) as a commie plot aimed directly at the heart of their beloved scholarly societies.

Another is less obvious, at least to me.... Faculty at small schools are unbelievably suspicious of their institutions' administrations. Anything that so much as

hints at institutional oversight of research gets the major fisheye. This makes IRs and institutional bibliographies a viciously hard sell.

A third: faculty everywhere are unreasonably paranoid about plagiarism, misquotation, and idea-appropriation, and unreasonably optimistic about copyright as a guard against those ills—but in my experience, faculty at small schools are worse, much worse. Open access isn't just a commie plot, it's a dagger in the heart of my precious ideas!...

I can imagine a small, tight-knit institution that really truly trusts its librarians making a go of an IR, especially if the technical pieces were handled outside the institution, by a vendor or consortium. It'd be a nice thing to see. I'm not especially sanguine, however; the barriers are significant, and by no means solely technical.

How can I comment knowledgeably? I've never even *worked* in a small academic institution (UC Berkeley hardly qualifies), and while my wife directed the library at such an institution, that was before IRs were a topic of discussion.

A December 1, 2008 post is entitled "IRs in 2009: the failure legacy." It covers a lot of ground, but I'm going to excerpt that portion most directly related to Green OA—and particularly the odd, damaging concept that Green OA doesn't really cost anything.

What is clear to me at this juncture is that the repository world is split in two as regards the appropriate response to faculty apathy about deposit. One chunk of the world is gamely gearing up to take on mediated deposit of the peer-reviewed literature. The other chunk of the world is using the repository for lower-hanging fruit (ETDs, undergraduate research, collapsing it with the local digital library, whatever) and doing a fan-dance around its lack of commitment to green open access. All right, there's a third chunk, too: those that are doing both.

However, it's the *second chunk* that are the problem, from the point of view of the open-access movement. Bluntly, these institutions *love* green open access—until it costs them resources beyond the mere provisioning of a repository. Since they are now uncomfortably aware that green open access costs more than that, they are sidling away from it in as delicate and face-maintaining a way as possible. The question for open access is how to keep its agenda alive in libraries that are no longer able or willing to dedicate specific resources to the green road.

There are options. One is to push harder on the gold road: ask libraries to dedicate funds to gold OA memberships and author fees. This is a tall order, because the subtext here is asking libraries to cancel even more journals to pay for a brighter future, but I do think gains are achievable, if perhaps mostly symbolic ones at this point.

Another is to bypass outreach to libraries and work harder on faculty, in hopes of additional Harvards and Stanfords... It's probably too soon to push this angle too hard in the States, unless targets are chosen very carefully indeed...

A third is to 'fess up about the real costs of green OA and provide libraries and struggling repositories with a realistic roadmap to achieving it. We have these data now; the problem is that libraries are backing away from the cost of the inevitable conclusions. If the open-access movement wants green OA through institutional repositories, it will have to stiffen the collective library backbone from the top down. I don't see this happening, but if green OA through IRs is to survive, it will have to...

Moving into 2009

In a couple of posts on March 2, 2009, Salo discusses Clifford Lynch's prediction that moves by funding agencies will mean "faculty will want help from their institutions in satisfying the requirements"—which, presumably, would be wonderful for IRs.

Salo's not biting quite that rapidly.

To which I say, great! *When it happens, let me know.* Libraries have already been sold one pig in a poke—that was the "faculty want someplace to put their papers" pig. I, as a repository-rat, have zero credibility left. Seriously, *none*. I can't go to my administration and say "Faculty are gonna want this. Really. Cliff Lynch said so!" That's a lead balloon, folks.

Wait until the NSF and NIH and Mellon get serious. *Then* go to libraries. Or better yet, tell *faculty* to go to their libraries; faculty asking for help have more credibility with library administrators than I will build up in my entire library career.

It gets worse. Lynch doesn't much care for OA as a basis for IRs. Quoting from a *Library Journal* article:

I may be a bit of an outlier on this, but it seems to me that IRs have become terrifically entangled with open access... I recently had cause to go back and reread my 2003 IR piece, and I hardly talk at all about OA, other than to observe that one of the many benefits of IRs is that they can facilitate OA mandates—but I consider them independent of one another...

Here's where things get strange—when Lynch questions conflating IRs with OA:

I think it is shortsighted. I know many of these institutions are feeling great pain from pressure on their acquisition budgets and would like to mitigate that... But that's a short-term economic thing, and I'm sorry to see it getting mixed up with IRs.

Whew. I was working full-time at UC Berkeley from 1968 through 1979. I can attest that, in the early 1970s, we had a large-scale serials cancellation project

because of pressure on the acquisitions budget. I suppose "short-term" could mean multiple decades, as opposed to centuries or millennia, but calling the serials affordability crisis "short-term" strikes me as odd. As it does Salo, commenting on that last quotation:

Oh, Dr. Lynch. Oh, Dr. Lynch. Mistake. Big one. The serials crisis is *not* a short-term economic thing; it's been here for thirty years and shows no sign of slowing. It's not a small thing, either; it's causing libraries significant pain... Pooh-poohing a major library crisis is not the way to win friends and influence people in librarianship, Dr. Lynch.

Anyway, if IRs aren't going to help solve the serials crisis, Dr. Lynch, why should a library invest in one? Riddle me that. "Altruism" is not a good answer here; we tried that one already. "Faculty demand" is an even worse answer, because *there isn't any*.

Moving elsewhere briefly, Richard Akerman provides another example of the challenges faced both by IRs and by OA in general, in his March 14, 2009 post, "revisiting potential research-support roles for the library." Key excerpts:

While institutional repositories are valuable, they currently benefit primarily organisations, not researchers. They provide a unified view of an organisation's published output. For individual researchers, their priority may be just on getting published, or if they do want to disseminate their work, they may just post it to their own website (and sad to say, may get more Google rank having it there than in their repository)...

It's clear that IRs have become about coercion, which should be making us seriously question their value...

If providing an institutional repository is your primary or core value to the organisation, you are putting yourself at tremendous risk, because a savvy administrator may notice that you can purchase hosted repository services from BePress and BMC Open Repository. Any time a primary function (however valuable) has become commodity, you are at risk.

What I do *not* see anywhere in the post (which, admittedly, is about other concerns): Any recognition that open access might itself *be* part of support for research. (Searching this blog for open access doesn't yield any real discussion more recent than mid-2007.)

The last stop on this hectic journey is an April 17, 2009 *Caveat lector* post, "Where the snark stops." Given the hard truths Salo has been stating about the past and present state of IRs—and the hard questions that need to be answered before IRs make sense—it's an important post. She speaks of a retired librarian's email referring to "the IR craze":

Whoa. Stop. Right. There. Please.

If I give the impression that I think IRs are a “craze” and should go away en masse, I hereby humbly apologize, because that has never been my belief or the intent of my communication. Do I think some individual IRs should go away? Sure. Do I think some of the assumptions underlying IRs need rethinking? Oh, of course. Do I hate IR software with a fiery passion? Yup, though I am gradually becoming friends with Fedora. Do I think some IRs are badly run? Absolutely, and I don’t except any of the ones I’ve been involved with, having no grounds whatever to let myself off any hooks.

But, see, I’ve been involved with them. For four years. I’ve staked my career on the damn things. I have invested time, money, sweat, political capital, professional capital, and a great many tears in them. I have, shall we say, a certain position from which to level criticisms.

Random librarians who email me just to sneer? Have no such position. If you don’t like your local IR, step up and *do something to make it better*. I’m guessing your local repo-rat would absolutely adore the help. If you can’t or won’t do that, don’t expect me to have much patience for your attitude; from where I’m sitting, you are part of my years-long headache.

It’s hard to comment on that. I read it and wonder again whether I’m adding value through these pieces or possibly otherwise. Does it damage OA for me to note (once again) that Harnad’s oversimplifications turn reasonable people away from OA? Maybe. Is it harmful to point out real-world issues, as Salo delights in doing? I don’t believe so (and, I would note, neither does Peter Suber). Is it *always* wrong to criticize something if you’re not ready to help make it better? Certainly not—particularly if you’re not in a *position* to improve things.

When I was, briefly, chair of the LITA Publications Committee, one of my four desires was to move LITA’s scholarly journal, *Information Technology and Libraries*, toward full and immediate Gold OA status for its refereed articles. (It’s already Gold OA with a six-month embargo for non-members, and it’s been Green OA for quite some time.) My success in that regard was similar to my success on the other three fronts: Zero. But I *did* try. Sniping without any effort to improve, when such efforts are available to you, is generally not such a great thing.

Can you have successful institutional repositories without green OA? Certainly. Can you have successful green OA without IRs? Probably not, particularly in the humanities, where disciplinary repositories seem thin on the ground. Can you run *useful* IRs (or disciplinary repositories) on a shoestring? I don’t believe so; Salo doesn’t believe so; and I believe it seriously

harms OA to assert that this is possible, or that Green OA is nothing but a matter of “a few keystrokes.”

Quality, Value and Progress

Peter Suber leads off this section with “Thinking about prestige, quality, and open access,” the lead essay in the September 2, 2008 *SPARC Open Access Newsletter*. (Note **2008**: it’s been around for a while.) It’s a rich, important essay. I’m providing *very* brief excerpts from a nearly 8,000-word essay. Specifically, I’m mostly just giving Suber’s theses with very little of the argumentation, and you should assume that all emphasis (*italics*, **bold**) is added, not in the original:

I’ve been thinking a lot lately about how journal quality and prestige overlap, how they diverge, and how their complex interaction affects the prospects for OA. Here are a dozen thoughts or theses about prestige and OA. Some are commonplace, but I include them because they help me build up to others which are not. I start with the rough notion that if journal quality is real excellence, then journal prestige is reputed excellence.

(1) Universities reward faculty who publish in high-prestige journals, and faculty are strongly motivated to do so. If universities wanted to create this incentive, they have succeeded.

(2) Most high-*prestige* journals today are toll access (TA). If a journal can be excellent from birth, but not prestigious from birth, or if new journals typically achieve quality before they achieve a reputation for quality, then we have a non-cynical reason to think that quality and prestige can diverge. Quality and prestige clearly overlap, perhaps most of the time. But a significant number of high-quality journals, most notably the new ones, will not be correspondingly high in prestige...

(3) Most authors will choose prestige over OA if they have to choose. Fortunately, they rarely have to choose. Unfortunately, few of them *know* that they rarely have to choose.

[To summarize: There are more and more high-prestige OA journals and most high-prestige toll-access (TA) journals allow Green OA.]

(4) Apart from the fact that most OA journals are new, there is no intrinsic reason why OA journals can’t be as high in quality and prestige as the best TA journals.

(5) Quality feeds prestige and prestige feeds quality.

(6) Prestige is a zero-sum game, but quality is not.

(7) Because prestige is a zero-sum game, and quality is not, prestige can actually interfere with quality.

When the journals in a field are few, it might be possible for the all the good ones to have recognizable brands and prestige in proportion to their quality. But when they are many, as today, then it’s difficult or im-

possible for all the good ones to have recognizable brands and prestige in proportion to their quality...

(8) Universities tend to use journal prestige and impact as surrogates for quality. The excuses for doing so are getting thin.

(9) Quality is made by authors, often in conjunction with editors and referees. Prestige is made by communities.

(10) Despite its value, prestige may only give TA journals limited protection against the rise of green OA.

(11) When OA journals approach TA journals in prestige, TA journals will lose their only remaining advantage. But this is *not* just a matter of time.

[A complicated and important argument that I'm unready to summarize.]

And the last few paragraphs. I'm not adding commentary because I don't have any—except to note that this is a cogent refutation of one-note, Green OA Only, advocacy.

University promotion and tenure committees should focus less on journal prestige and journal impact than on article quality and candidate quality. I know that's easier said than done. We'll never have quality metrics that are as easy to apply as our current prestige and impact metrics. But we can stop putting easy judgments of prestige or impact ahead of difficult judgments of quality, and we can find help in metrics which oversimplify less than the one we tend to use now.

When prestige and quality diverge, journals, universities, and authors all tend to favor prestige. It's not hard to see why. When prestige and quality diverge, prestige continues to offer undiminished rewards and create undiminished incentives. Quality is a weaker incentive when it is not accompanied by prestige. Journals have their own reasons for favoring prestige over quality: because of the zero-sum problem, prestige boosts quality more than quality boosts prestige. But authors favor prestige mostly because their universities lead them to, and universities tend to favor prestige because it's easier than favoring quality. If universities could take on the difficult job of assessing quality, they'd change incentives for authors, which would have at least some effect on journals.

Prestige is a real incentive, for journals, universities, and authors. We shouldn't expect that any of these players will nobly rise above prestige. But neither should we underestimate the attraction of prestige or its superior attraction when prestige and quality diverge. Nor should we underestimate either its non-accidental relationship with quality or the non-cynical reasons for thinking it can diverge from quality. Nor, finally, should we underestimate either side in a delicate balance of opposites: our own role in creating prestige and the difficulty of creating prestige where it doesn't already exist.

Prestige is no obstacle to green OA. But green OA suffers when authors make the mistaken assumption that publication in a prestigious TA journal is incompatible with OA. Prestige is a greater obstacle to gold OA, but only because of gold OA journals are new, not because they are OA.

Two developments would change everything: (1) roughly equal prestige for OA and TA journals of roughly equal quality, regardless of age, and (2) high-volume green OA across the disciplines. (Funder and university OA mandates are terribly important, but they are merely means to the second of these.) The two developments are compatible, and we can work for both at once. We can make rapid progress on the second as soon as we have the will. But we can't make rapid progress on the first, even with the will, and my main purpose in this article has been to show why. We can describe the impediment from many angles: the benign circle entrenching the high-prestige TA journals, the vicious circle excluding newer OA journals, the zero-sum game of prestige, the slow-changing community attitudes that create prestige, the slow-changing allocation of funds paying for peer-reviewed research articles, and the stubborn fact of age. This impediment doesn't prevent OA journals from becoming first-rate, or even from growing in prestige, but it slows progress, like the slope of a hill, and can deprive OA journals of the feedback effects which boost submissions, quality, and prestige.

The second development is attainable, as advertised. But the first is equivalent to the state in which quality and prestige never diverge, which shows that it's an asymptote. We can increase the prestige of some OA journals, and sometimes even bring their prestige into alignment with their quality, and the same is true of publisher efforts on behalf of new TA journals. But we'll never prevent quality and prestige from diverging. In my mind, these are reasons to work for gold OA and green OA simultaneously: gold OA, so that we don't further delay the benefits of hard-won, slow-growing incremental progress, and green OA so that we don't miss precious, present opportunities for accelerating progress.

Crossover points toward universal open access

That's my title, but I think it's a fair one—for the second portion of Peter Suber's "Predictions for 2009" in the December 2, 2008 *SPARC Open Access Newsletter*. (I'm not including the first because it's a little late in 2009 to be noting predictions—and too early to compare predictions to reality.)

Suber lists a dozen crossover points in the order in which he predicts they will occur. I'm providing the crossover points themselves and relatively small portions of Suber's explanations—and it's probably

important to repeat that the order is *not* importance, but predicted reality, from soonest (in the first two cases, it's already happened) to last. It's probably also important to note that Suber defines crossover as passing the 50% mark. He says a true crossover point may be a much lower percentage in some cases; I'll suggest that it might also be much *higher* in some cases. Again, added emphasis is my own, not Suber's.

One finish line is **total success**: immediate OA to 100% of new research literature. But I don't include that kind of goal here, partly because it's not a crossover point (we'll probably approach 100% OA for new articles asymptotically) and partly because I want to focus here on strategy, which could always be clearer, rather than the ultimate goal, which is already clear. For strategy, we have to change a set of defaults. If today most new research articles are never self-archived, then flipping that over so that most new articles are self-archived will be a critical, game-changing milestone. That's what I call a crossover point.

Herewith, the dozen predictions, with brief excerpts [or my quick summary in square brackets] of Suber's commentary. (In this case, the portion of the article is less than 3,000 words—and it's a quick read.) My comments, if any, appear in unindented type.

(1) **Green OA permission.** Today about 63% of surveyed TA publishers, and virtually all OA publishers, allow authors to deposit their peer-reviewed manuscripts in OA repositories. We reached the cross-over point years ago when a majority of new articles carried permission for green OA.

Even though we've crossed this threshold, and today many more surveyed publishers allow self-archiving than not, the percentage has stopped growing and started declining. It may be that some journals are retracting their permission. Or the decline may just be just a side-effect of the order in which SHERPA is surveying journals, and it's turning out that there are fewer green journals among newly-surveyed titles than among those surveyed earlier. I'm not worried, however, that we'll ever cross *back* over the cross-over point...

(2) **OA books.** Even the youngest scholars today grew up in a world in which there were more print books in the average university library than gratis OA books online. But that ratio flipped over about two years ago. Today there are many more gratis OA books online than print books in the average academic library, and we're steaming toward the next cross-over point when there will be many more gratis OA books online than print books in the world's largest libraries, academic or not. The two analogous cross-over points for libre OA books are much further off...

I'd question the significance of this milestone, given that the "gratis OA books" are almost entirely public

domain, thus either government publications or (generally) pre-1923. I might even question the numbers (are there really as many public domain books available for scanning projects as there are books in the Library of Congress or Harvard or even Berkeley?), but don't know enough to make that an intelligent challenge.

Everything below this is *somewhere* in the future.

(3) **Funder policies.** Today most publicly-funded research is not subject to OA mandates. At the cross-over point, most of it will be. Because public funding agencies differ enormously in size, even within the same country, we'll reach the cross-over point long before most public funding agencies adopt OA policies. A critical set of the big ones will get us there, and we already have many of the biggest, including the NIH, six of the seven Research Councils UK, and a pilot OA mandate from the EU's FP7...

(4) **Green OA deposits.** Today only a minority of new peer-reviewed manuscripts are self-archived soon—say, within a week—after they are accepted for publication. The cross-over point will come when most are self-archived soon after publication. In certain fields, like particle physics, the cross-over point is behind us. But for the total corpus of new journal literature it's still well ahead of us...

(5) **Author understanding.** Today only a minority of publishing researchers has an accurate understanding of OA. The cross-over point will come when that minority becomes a majority. Author understanding is undoubtedly improving, but the slope of the curve is shallow rather than steep. Why is it shallow, and why didn't we reach this cross-over point years ago? We're up against myths and misinformation from publishers. We're also up against misunderstandings from excited newcomers. We're up against a culture of university research in which researchers who are unfamiliar with OA are preoccupied with their research, overworked, uncritically focused on the publishing incentives created by their promotion and tenure committees, and grateful to sign just about any contract a journal puts in front of them.

(6) **University repositories.** Today most colleges and universities lack institutional repositories. The cross-over point will come when most have them, either individually or as members of a consortium. Why will this point come after the funder-policy and green-deposit cross-over points (#3 and #4)? First, some funders use their own repositories, the way the NIH uses PubMed Central. Second, some faculty will prefer to deposit in disciplinary repositories, like arXiv, even when they have an institutional repository. Third, universities move more slowly than funding agencies (see #10), even if they have the same interests as funding agencies in advancing OA.

Fourth, IRs have been decidedly problematic, as Dorothea Salo's been telling us—and those problems probably aren't helped by repeated claims that IRs are essentially free. Fifth, trying to restrict IRs to the published scholarly literature makes it even more difficult to sell them to faculty.

(7) **Libre gold OA.** It's hard to be sure, but it seems that most OA journals are merely gratis OA today. While they remove price barriers, the large group I'm talking about removes no permission barriers and leaves users with nothing more than fair use or the local equivalent. The cross-over point will come when most OA journals will be libre OA, and remove at least some permission barriers.

Frankly, if the definition of libre only requires removing *some* permission barriers, I'd expect this point to be reached fairly rapidly. Maybe there need to be goalposts: e.g., a majority of gold OA publishers having the equivalent of Creative Commons' BY-NC license would be one (and should be reachable fairly easily, I'd think), where getting a majority to the BY equivalent would be considerably more difficult—and getting them to remove *all* permission barriers (the equivalent to CC 0, as far as I can tell), may be nearly impossible...

(8) **Journal backfiles.** Today most TA journals don't have OA backfiles. At the cross-over point, most will. The cost can be large, and we won't see significant progress until some of the wealthy book-scanning projects move more aggressively into journal-scanning...

(9) **Author addenda.** Today only a small number of universities have adopted author addenda, and only a smaller number require their use. That makes it easy for publishers to reject the addenda or to refuse to publish work by faculty for whom addenda are not optional...

(10) **University policies.** Today most university research is not subject to university-level OA mandates. At the cross-over point most of it will be. Although 24 universities and four departments already have OA mandates, a total which is very close to the number of funding agencies with OA mandates (30), I'm predicting that the cross-over point for universities will come much later than the cross-over point for funders.

A mild prediction: The whole scheme of universal mandates will prove so unsavory that it will never reach anywhere close to 100% success, at least in the United States and among liberal arts institutions. If mandates are the only way we're going to get OA, then OA is in serious, serious trouble.

(11) **OA journals.** Today most peer-reviewed journals are TA. At the cross-over point most will be OA. Two kinds of events move us toward this point: the launch of new OA journals and the conversion of TA journals

to OA. And two kinds of event move us the other way: the launch of new TA journals and the conversion of OA journals to TA. All four events occur, although the first three occur much more often than the fourth. On the whole we're moving toward the cross-over point, rather than away from it, but the movement is slow...

Suber correctly (in my opinion) notes that the cross-over may be slow because "some TA journals can survive in an OA world"—either because they're highly prestigious, because many or most of their subscriptions are *not* from academic libraries, or because they charge so little that there's no solid economic reason for conversion.

Of course, we *could* reach this crossover point by having dozens of new OA factories, each churning out hundreds of new OA journals. But the track record for solid peer review and high quality standards of recent OA factories—that is, those that announce scores of new titles at once—is so mixed (that's not the word I want to use) that this may be a recipe for disaster, not success.

For that matter, counting journals becomes silly at some point. If 50% of all supposedly peer reviewed journals are Gold OA, but those journals publish only 10% of all refereed papers and only 5% of the quality papers, no useful crossover point would have been reached, other than perhaps the point of madness for anyone attempting to keep up with the literature.

(12) **Libre green OA.** Today most green OA is gratis OA. But at least some is libre OA, for example, a subset of PubMed Central. There may be similar subsets in many other repositories. As long as most green OA depends on the permission of TA publishers, the majority of it will remain gratis...

Lots more discussion in the original article.

Is OA anti-publishing

By now you've doubtless figured out that *scholarly kitchen* is a hotbed of anti-OA blogging (among other things). Kent Anderson's December 2, 2008 post, "Are publishers anti-publishing?" further demonstrates that. The post itself is interesting and makes some valid points, but here's the money paragraph (emphases added):

Investment and change are tough. For STM publishers, it's just as difficult to avoid the anti-publishing trends (devaluing content, missing the message from users/subscribers). **The Open Access movement is implicitly anti-publishing—essentially, the value proposition is that content is so worthless that you have to pay to have it published, and you can't charge for it even then.** That's anti-publishing, treating content as less than a commodity. Institutional site licenses drop the paid model to cents on the dollar, undervaluing published works by depre-

cating them to a cost basis rather than recognizing their value to academic pursuits. **Institutional repositories are anti-publishing, not trying to reach a broader audience but trying to showcase their institution's "intellectual property" for a purpose that is implicitly anti-publishing.**

Let's see: 75% of toll-access journals charge author-side fees. More than 70% of OA journals do not. So, by Anderson's "reasoning," it's toll-access publishers that are anti-publishing. The second bolded item is interesting in that it *explicitly* denies that IRs create broader audiences for scholarly articles. I don't know about "implicitly anti-publishing," but this is explicitly nonsense.

Kevin Smith has another and maybe more appropriate response, as cited by Peter Suber, in a December 5, 2008 item at *Scholarly Communications @ Duke*, "What is 'value' in publishing?":

Which brings me to where Anderson really goes wrong—his comments about how open access and institutional repositories are "anti-publishing." To get to this claim one must define publishing very narrowly, based on a traditional, "the way we have done it in the past," standard; Anderson sounds a lot like Gleik at this point. On-line, open access distribution IS publishing, of course, as the many peer-reviewed open access journals clearly prove. What is most astonishing about Anderson's discussion of these "anti-publishing" trends, however, is his claim that open access "devalues" scholarly content by "treating it as less than a commodity." How can one make such a claim about scholarly content when authors have been expected to give their writings away for free to publishers for many years? Scholarly authors are used to thinking about the value of their work in terms other than economic, and those terms have been dictated, in part, by the business model of traditional scholarly publishing.

The value of scholarly work, for scholars, has never been based on the money it could earn, since they never saw a penny of that money and were, in fact, expected to pay for access to their own writings. Often they were even expected to pay "page charges," which makes the author-side fees now charged by many publishers for open access seem very familiar. The point is that access and use, not economic gain, define the value of scholarly writing because they serve the scholarly authors' need for recognition and impact; the cost of the wrapper in which the work was contained (the commodity) has never been a marker for value in the academic world, and it has lately become an impediment.

I fervently hope that scholarly publishers can find ways to add value to academic content, as Anderson challenges them to do. But that task will be much more difficult if it is based on a narrow view of the value of academic work that begins and ends with the

traditional way publishers have done business. The search for new models of scholarly publishing will have to take into account the things that actually matter to academic authors and scholarly institutions.

Other than specious attacks against OA, what Anderson's stuff may boil down to is similar to what I've seen in complaints that "books are dying" because the big New York publishing houses may be losing ground to the fleet-footed small publishers and self-publishers: That is, Anderson implicitly means "Publishing with a capital P"—the kind of Traditional Publishing that *requires* big chunks of capital and expects to earn big returns on that investment.

Abridgment as added value

There are quite a few ways that Gold OA publishers can legitimately charge for added value, and with luck those ways could work well enough to avoid the need for author-side fees and still yield sustainable publishing. (I won't say *profitable* because there's no reason university OA publishing, library OA publishing or society OA publishing needs to do more than recover costs—if that. There's nothing inherently evil about profit, but there's nothing inherently sacred about it either.)

So, for example, a journal that combines peer-reviewed papers with other material—book and website reviews, editorials, news briefs, etc—can charge for an electronic or print version that includes the other material. A journal that's e-only can be available, for a price, in an annual print volume *without any capital investment* (if the articles are already in formatted form), using print-on-demand techniques.

Peter Suber discusses another added value in the lead essay in the September 2, 2009 *SPARC Open Access Newsletter*. Excerpts:

Imagine that an open access (OA) journal could generate revenue by selling abridgments of its full-text OA articles. Imagine that the revenue even made it unnecessary to charge author-side publication fees. That would be a supremely elegant business model, if only it could be made to work.

BMJ has made it work for more than 10 years, and next year will take the idea even further.

All BMJ research articles are full-text OA in the digital edition of the journal. The print edition, which is toll access (TA), contains 3-5 page abridgments of each research article. BMJ calls this system ELPS (for "electronic long, paper short"). The OA edition of the journal charges no publication fees, and the full-text research articles have no word limit.

Nine months ago, BMJ introduced even shorter, one-page abridgments called picos, and published selected articles in the pico format rather than the long-

er 3-5 page format. The experiment has been so successful that BMJ announced last month that it will phase out ELPS and go all-pico. Starting in January 2010, all BMJ research articles will have pico abridgments in the TA print edition, and the full texts will still be no-fee OA in the digital edition...

[Because authors will be doing the pico versions and BMJ editors were writing the longer summaries, this *directly* saves money for BMJ. A survey also suggests that authors and readers both prefer picos.]

Picos will occupy about 8 of the 60 pages, or 13%, of each print issue of BMJ. The rest will be devoted—in Groves' words—to “education, clinical reviews, analysis, news, other medical journalism, letters, and views and reviews.”

That's the model. Here are a few thoughts.

- The BMJ survey results don't surprise me at all. I'd love to see structured one-page summaries of new articles in my field...
- This model works better for some kinds of summaries than others. As BMJ found, pico-length summaries have more marketing power than shorter abstracts and longer ELPS summaries. The sweet spot will probably differ from field to field.
- Moreover, if abridgments are done badly, they'll provide no added value at all, or the subtracted value will exceed the added value...
- If a journal's summaries hit my sweet spot for length and quality, would I be willing to pay a reasonable price for them? Absolutely...
- I'd find it even easier to pay for TA summaries if the full texts were OA...
- Boiling down a long article to a one-page summary could give us value we didn't have in the original. Authors know this if they've ever written an abstract and found that the exercise gave them new clarity about what they were saying...
- While few OA journals have TA editions, many TA journals have partial OA editions. Those TA journals might consider picos as part of a larger business model allowing them to provide OA to their full-text research articles...
- OA journals without TA editions could consider partial TA editions. This wouldn't detract from OA if all their research articles were still full-text OA...
- Bottom line: Summaries appeal to a much larger audience than full-text articles. The market is larger. The demand is greater. That's a fact worth knowing for any journal wishing to generate revenue by selling some content and providing OA to the rest.
- Picos don't merely appeal to a larger market than full-text articles. They better fit the economy of attention. When we need to read full-text articles, we need to find the time to do so, and we often

fail. But whether we need to read full texts or merely skim, summaries are friendlier to our crowded, demanding lives. We're ready to pay for services that save us time...

- [Several more worthy points omitted]
- I'm not taking a position on whether summaries ought to be OA or TA. I'm saying that selling access to summaries in order to support full-text OA is elegant and promising, even if counter-intuitive, and should be considered in every niche where it might work. If intelligent summaries really benefit readers, authors, journals, and OA, even when the summaries are TA, then they are an experiment worth testing and replicating elsewhere.

Open Access in 2008

Way too long to include or even make a serious stab at excerpting (the *JEP* version is just under 13,000 words—half of this interminable roundup!), but way too important to ignore, this is Peter Suber's roundup of last year's trends. You'll find one version in the January 2, 2009 *SPARC Open Access Newsletter* and another in the (ahem) Gold OA ejournal, *The Journal of Electronic Publishing* (*JEP*), vol. 12, no. 1 (February 2009), the special OA issue.

Here, just for fun, are Suber's “Highlights of the highlights”—“top 10” lists of the worst and best OA events of 2008 (if acronyms and initialisms seem mysterious, read the article):

The worst of 2008:

10. The Elsevier rules for electronic interlibrary loan: make a printout, scan the printout, and loan the scan. Not unique to Elsevier, but still a kludge to avoid taking advantage of 20th century technology.
9. The State of Oregon's claim of copyright in its statutes. An indefensible claim leading to groundless legal threats. Relinquished when challenged.
8. The American Psychological Association's \$2,500 fee for depositing author manuscripts in PubMed Central. An attempt to charge for green open access as if it were gold open access. Withdrawn when publicized.
7. The UK Ordnance Survey use of public funds to pay a lobbying firm to push back against public pressure to provide open access to its publicly funded mapping data. It was bad enough to charge taxpayers a second fee for access to publicly funded data. Now it charges them for the lobbyists who work against their interests.
6. The 35-year embargo adopted by the American Anthropological Association for the OA backfiles of its two leading journals. The longest embargo or moving wall I've ever seen a publisher boast about. More than 10 times longer than the runner-up.
5. The Google settlement's effects on fair use. The settlement has many compensatory gains, but Google

dropped a winning case and its willingness to pay for fair-use copying may force others to pay as well.

4. The 12-month embargo allowed by the NIH policy. Twice as long as the embargo used by any OA-mandating medical research funder in the world. A needless sop to the publishing lobby, and one that slows down medical research.

3. The American Association of University Presses support for the Conyers bill (a.k.a. Fair Copyright in Research Works Act). The presses said they were not opposing the NIH policy but so far have not refined their support for the bill to target the practices they dislike and spare the NIH policy.

2. The continuing slow pace of spontaneous self-archiving. Still due to inertia more than opposition, but still a problem. Habits die hard, especially when the proposed alternative is unfamiliar, widely misunderstood, and orthogonal to entrenched incentives.

1. The Fair Copyright in Research Works Act. Amending copyright law to block an OA policy consistent with current law, while pretending to be motivated by the policy's copyright violations. Harmful bill + misleading title + deceptive rhetoric, brought to you by lobbyists paid with your subscription dollars.

It's not hard to list obstacles and setbacks to open access. We encounter them every day. But a Top 10 list may leave the false impression that some of the obstacles are insuperable. Even the #1 item on this list, however, doesn't merit that description. The bill didn't come up for a vote in the last Congress, and the September hearing on it before the House Judiciary Committee educated many committee members who previously knew nothing about the NIH policy or scholarly publishing but what the publishing lobby had told them. Some didn't know that 25+ public funding agencies in other countries had already adopted OA mandates or that all the mandates covering medical research used embargoes half as long as the NIH embargo. Some didn't know that scholarly journals don't pay authors or peer reviewers. Some didn't know that the NIH spends about \$100 million/year to support the publishing costs of private-sector journals, or about 30 times more than it spends implementing its public-access policy. The bill may come back in the next Congress, but it's a transparent giveaway to special interests at the expense of the public interest and we can defeat it.

The best of 2008:

10. The launch of Europeana. Already large and valuable, on track to be more so, unifying many smaller projects, and committed to open access for its public-domain contents.

9. The launch of OAPEN and Bloomsbury Academic. Two notable examples, among many others, of book publishers committing themselves to the synergies of

open access and print on demand for academic monographs.

8. The launch of the Open Access Scholarly Publishers Association and the SPARC Europe seal program. Two bottom-up ways to set standards for the fast-growing world of OA journals.

7. The libre OA mandates from the UKPMC Funders. A first for funding agencies: demanding the removal of permission barriers, not just price barriers, when funders pay part of an article's publication costs.

6. The profitability of BioMed Central. Significant new proof that gold OA is sustainable. Disarming one more objection and removing one more excuse.

5. The European Research Council OA mandate. The first and strongest EU-wide OA mandate, covering 15% of the FP7 research budget.

4. The EU OA mandate and pilot project. A strong policy for another 20% of the FP7 research budget, and a sign of what's to come for the rest of the research budget.

3. The promise of the Obama administration. Open access supporters in key agencies and in the new President's Council of Advisors on Science and Technology. A sea change from the Bush administration, which forthrightly subordinated science policy to corporate interests.

2. The Harvard OA mandate and the Harvard effect. A first-rate policy from a first-rate university, backed by a unanimous faculty vote, inspiring change and proposals for change at other universities.

1. The NIH OA mandate. As I put it at the time: "Measured by the ferocity of opposition overcome and the volume of literature liberated, it's the largest victory in the history of the OA movement." If the policy isn't repealed, this could remain the case for years and years.

Miscellany

Herewith, a few more items that don't neatly fit together, except that they're all concerned with OA. I'll take them in chronological order—*except* for the last one, which is actually the oldest of the lot.

Flat World Knowledge and U.S. Gov't on Open Access Course Materials

Peter Murray, June 5, 2009, *Disruptive library technology jester* (dltj.org/). Briefly, it's about two things that *could* lead to some system of open-access textbooks and related academic material.

The first is Flat World Knowledge, a startup with \$8 million in VC funding. Excerpts from that portion of the post:

Flat World Knowledge launched earlier this year, and it brings an entrepreneurial feel to the staid subject of

textbooks. Billed as “the world’s first publisher of open-source college textbooks,” their website has a scrappy, web2.0 start-up feel to it... To faculty and staff in higher education, Flat World Knowledge describes themselves this way:

We preserve the best of the old – textbooks by leading experts.

Then we flip it on its head.

Our books cost \$0 online. We provide paperbacks, audio books, and self-print versions for under \$30. Our books are open for you to edit for your class. Our new editions are on your terms. We publish them – you decide if and when to use them.

They offer free versions of their textbooks online then charge for various derivatives and additions. Instructors can modify the textbook — rearranging chapters, add or delete chunks of text, and (coming soon according to the site) be able to add materials based on a database of what is available at Flat World Knowledge. (One has to register on the site to do this, but you can watch a video tutorial to get an idea about how it works.)

[Murray lists the textbooks currently available and those in the pipeline.]

According to both commenters, the trouble with this concept is that at least some of the initial offerings aren’t very good. One commenter calls some of them “well-known ‘road-kill’: i.e., they died an ignominious death in the traditional textbook market, and then the rights to them reverted to the author” and notes that none of them seems to be well-edited. The other commenter tags one book as a reread of a failed textbook and questions the quality of all the content. It would be inauspicious at best if the first serious attempt at open access textbooks consisted of dreck.

The second part discusses HR1464, a bill that—like all good bills these days—was apparently named by that secret agency in charge of cute acronyms: the Learning Opportunities With Creation of Open Source Textbooks (LOW COST) Act. (Yes, “open source” should be “open access,” as Murray notes, but LOW COAT isn’t nearly as cute.) The bill would require federal agencies spending more than \$10 million on science education to spend 2% of those budgets on the development of college-level educational resources. It’s a start—maybe. (As Murray notes, there were no cosponsors for the bill at introduction, never a good sign. As of September 26, 2009, there are *still* no cosponsors.)

Adventures in web publishing

This one’s too interesting to pass by—but it’s also one where I think you’re better off reading the original, by

Douglas J. Amy in the June 18, 2009 *Inside Higher Ed* (www.insidehighered.com/views/2009/06/18/amy). It’s brief (1,380 words) and well written.

Amy—a tenured full professor with three published scholarly books to his credit—wrote a manuscript intended for a general audience, *Government is Good: An Unapologetic Defense of a Vital Institution*. After approaching publishers, he didn’t believe that the book could get the kind of exposure he wanted. So he put it on the web—on a properly-designed site, not just as a PDF. (He got a grant for \$2,000 to have someone design and implement the website.)

Results? More than 75,000 people have visited the site, half of them apparently staying long enough to *read* parts of the book. He’s satisfied that he’s reached a broader readership than for all of his other books, and there have been other good consequences. It’s not an experiment that would suit everyone, but it’s an interesting approach.

“Don’t ask, don’t tell” rights retention for scholarly articles

Stuart Shieber posted this on June 18, 2009 at *The Occasional Pamphlet*. Excerpts:

A strange social contract has arisen in the scholarly publishing field, a kind of “don’t ask, don’t tell” approach to online distribution of articles by authors. Publishers officially forbid online distribution, authors do it anyway without telling the publishers, and publishers don’t ask them to stop even though it violates contractual obligations. What happens when you refuse to play that game?...

Shieber notes the reasons to publish research, the traditional means, why scholars are motivated to publish and the need both to publish through a peer-reviewed journal and to make research widely available, which the web itself does very well.

An author has a simple solution to the quandary of whether to distribute through a publisher’s access-limited mechanism or freely online: *Do both*. Unfortunately, publishers typically restrict authors from this approach through contractual limitations stipulated in copyright assignment forms.

This brings us to the strange social contract. What has arisen, perhaps surprisingly, is a kind of “don’t ask, don’t tell” approach to online distribution by authors. Publishers officially forbid online distribution, authors do it anyway without telling the publishers, and publishers don’t ask them to stop even though it violates contractual obligations.

The standard system for scholarly communication is thus based on widespread contractual violation and fraud.

After noting that publishers *could* police violations much as the RIAA does—and that a few take-down letters would probably have a chilling effect on scholars who post their papers.

We can only speculate that the fear of upsetting their content providers trumps their need to maintain control over the content itself, given that there is no evidence that the online availability is hurting their revenues... One can just imagine what kind of backlash an RIAA-style approach would have with academics, in addition to the desired chilling effect.

Shieber is not willing to routinely violate contracts. Instead, he amends the conditions of copyright assignments and will not sign an assignment form that doesn't allow for online noncommercial distribution. In general, publishers haven't complained—but there's one exception.

The remainder of the post details that exception. It's an odd story, one that I can't excerpt very well. It boils down to one Blackwell philosophy journal *absolutely insisting* on a 12-month embargo before Shieber could post his own article online—and pulling the article when he would not agree. (Shieber details three arguments Blackwell's editor gave for this insistence, all of which he regards as specious.)

Shieber pushed the matter to the journal's editorial board—and, eventually, Blackwell changed the policy, abandoning the required embargo for this particular journal. The article was published, a mere three years late.

What is the moral of this story? First, all participants—including the editorial board and editor-in-chief of the journal, the managing editor, the publisher's staff—were people of good will. They all acted in ways they thought in the best interest of the institutions they represented and the larger missions of those institutions. However, to a great extent, they may not have fully thought out the connections between the policies they acted under and the missions. The editorial board may not have realized that the journal's policy embargoed author distribution; certainly the journal's contributors didn't, or chose to ignore it. The publisher may not have realized the inconsistencies between the journal policies and the facts-on-the-ground.

But it is also apparent that authors are far too acquiescent in the process of rights retention with publishers. We are overly willing to accept the rulings of publishers as a *fait accompli*. Despite the fact that publishers assert that their policies are supported by their editorial boards, editorial boards are in fact responsive to reasoned arguments. And although a negotiation for rights retention between an author and a large commercial publishing company asymmetrically

disfavors the author, one in which the author is supported by the editorial board is a different matter entirely. This example calls for taking advantage of rights retention negotiations to enlist editors and editorial boards in the process of expanding access to scholarly articles in a way consonant with law, moving past the “don't ask, don't tell” social contract.

Well said.

A field guide to misunderstandings about open access

This first appeared in the April 2009 *SPARC Open Access Newsletter*, but it also appears as a separate item at SPARC: www.arl.org/sparc/publications/articles/openaccess_fieldguide.shtml. A considerably shorter version—but still more than 5,000 words—appears on the Library Leadership Network as “Open access misunderstandings.”

The 10,000-word field guide cites and discusses 25 common misunderstandings, how they arise and more. **Go read it**—even if all the “misunderstandings” may not be entirely false.

Conclusion

That's the gamut—a tiny fraction of what's been said about OA in the past year, but *much* more space (if with a much different excerpt-to-commentary ratio) than I'd normally devote to OA in an issue or a year.

Is this a coherent overview of OA, or OA between mid-2008 and now? *Absolutely not*. For that, you'd be better off with Peter Suber's overviews or with the Open Access cluster in the Library Leadership Network (which I highly recommend, having written or edited most of it): Start at the Category page (pln.lyrasis.org/wiki/index.php/Category:Open_access) and go from there—starting with “Open access basics” and preferably reading the whole set.

The question now is whether LIBRARY ACCESS TO SCHOLARSHIP should or will remain as an occasional feature in *Cites & Insights*. Here's what I had to say about it on *Walt at Random* (with modifications):

Why I'm considering dropping the section

Value added: I've never felt I could add much value to Peter Suber's commentaries or, for that matter, Dorothea Salo's (when she was focusing on these issues). I've given up engaging Stevan Harnad or directly discussing his monotone writing. Lately, I'm not sure my synthesis and commentary are adding much value to any of this.

Effectiveness: Most *Cites & Insights* readers are within the library field, I believe—and that's only reasonable, since that's my background and the focus of most topical areas. So I'm probably not reaching many scientists—or, if I am, I'm probably not doing much to convince them to do

more about OA and access-related issues. As for librarians, I'd guess that my readers are mostly already convinced—that I'm neither educating nor convincing much of anybody who doesn't already get it. (I'd guess 1% to 3% of librarians read *C&I*, spiking to 25% or more for one particular issue. Those who need educating are mostly in the other 97%, I suspect.)

Futility: Given what I'm reading from scientists as to how they relate to libraries and librarians, and given what I'm reading as to how they make decisions on where to publish and where to exert pressure, I'm feeling pretty futile about the whole effort. Not **necessarily** about OA **as such**—but definitely about my ability to make a difference.

LIBRARY ACCESS TO SCHOLARSHIP essays appear to be read and downloaded a *lot* less often than essays on blogs and blogging, Google Books, wikis and the like and somewhat less than essays on copyright and MAKING IT WORK.

More reasons for abandoning this section, reasons that admittedly overlap with the three above:

- The addition of Bill Hooker's *Open Reading Frame* and Stuart Sheiber's *The Occasional Pamphlet* may make my contributions even more superfluous.
- It's difficult to escape the conclusion that the "OA community"—the bloggers who focus on open access, notably apart from Peter Suber and Charles W. Bailey, Jr.—would be just as happy if I disappeared or, perhaps more correctly, have never been aware (or cared) that *C&I* even existed.
- I grow increasingly convinced that most scientists *just don't care*—either about libraries or about OA—and maybe that's appropriate. I also grow increasingly convinced that librarians can't do it on their own, although it's encouraging to see things like the Compact that recently emerged. Still, it's an uphill battle, and one that I really can't play much part in.
- Every time I see calls for "universal mandates," I want to back as far away as possible.
- **One new one:** Sometimes it seems as though it's all been said, that we're now engaged in endless rehashing.

On the other hand...

I came up with these reasons in that second post:

- Maybe, just maybe, I'm occasionally putting points together in a different manner.
- Maybe, just maybe, I'm persuading one or two people a year to think about aspects of library

access to scholarship they might not have thought about otherwise.

As I noted in the blog, the reasons to keep going don't look like much of a hand; I'm inclined to fold.

For now...

I've done this roundup. I intend to be *very* cautious about tagging any new items. After a year or so, I'll either find that, in the few cases where I can't avoid them, OA items fit elsewhere (e.g., TRENDS & QUICK TAKES) or that I feel the need to restart this section.

If you want to keep track of OA but don't want to spend much time on it, a couple of suggestions:

- Subscribe to the *SPARC Open Access Newsletter*. It's free, it arrives once a month and, while it's long, *you don't have to read the whole thing*. Read the title and first paragraph of Suber's lead essay to see whether it's a topic you want to pursue; then skim through some of the other material, if you're so inclined.
- Subscribe to *Open Access News* through an aggregator—and skim through the titles quickly. Yes, there's a lot here; no, you don't have to read all of it. I mention OAN first among the blogs because it's by far the most balanced and most comprehensive.
- Try one or two other OA blogs. Some are cheerleaderish; some are too simplistic; some deal with fairly sophisticated issues; several have much to offer.

Will we ever reach full OA? Would that be a good thing? For now, I choose not to speculate.

Masthead

Cites & Insights: Crawford at Large, Volume 9, Number 12, Whole Issue 122, ISSN 1534-0937, a journal of libraries, policy, technology and media, is written and produced by Walt Crawford, Editorial Director of the Library Leadership Network.



Cites & Insights is sponsored by YBP Library Services, <http://www.ybp.com>.

Opinions herein may not represent those of Yraxis or YBP Library Services.

Comments should be sent to waltcrawford@gmail.com. *Cites & Insights: Crawford at Large* is copyright © 2009 by Walt Crawford: Some rights reserved.

All original material in this work is licensed under the Creative Commons Attribution-NonCommercial License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/1.0> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

URL: citesandinsights.info/civ9i12.pdf