**Academic Libraries as Scholarly Publishers**

Dorothea Salo  
*Digital Repository Services Librarian*  
*George Mason University*

“I amar prestar aen: han mathon ne nen, han mathon ne chae, a han noston ned

‘wilith.”

JRR Tolkien, *The Return of the King*, translated from English by David Salo

**Introduction [A]**

Academic librarians have historically been the scholarly literature’s collectors and preservers, its major purchasers and ultimate disseminators. What happens between author manuscript and ink (or, increasingly, pixels) on journal pages has rarely been our concern. The world is changing, however; to maintain acceptable levels of access in the face of unendurable serials price increases and shifts in how researchers communicate results to each other, we are moving (willingly or not) toward more and earlier intervention in the scholarly-publishing process.

**Fuel for Change [A]**

University presses and scholarly societies managed functional, affordable publishing models for quite some time. What happened?

Many scholarly societies, especially in the hard sciences and medicine, sold their journals outright to large for-profit publishers who contributed back a cut of the proceeds; others came to depend on journal subscription revenue to fund other society activities. For-profit publishers grew, merged, and came to control large swathes of the scholarly record, allowing them to squeeze profit from low-readership, low-profit journals by bundling them with crucially
necessary ones. The resulting “big deal”\(^1\) has broken library budgets, even consortium budgets; ripple effects have decimated the ranks of university presses, causing publication difficulties for humanities and social-science scholars whose tenure and promotion rests on the publication of scholarly monographs.

Times have not been easy for publishers, either. The advent of electronic publishing and distribution meant expensive experimentation and workflow retooling for publishers who suddenly had to produce two or more versions of journal content. Keeping electronic content from unauthorized eyes required additional investment in staff and technology. The sheer amount of research skyrocketed, causing especial problems for well-read, high-impact journals whose submissions overwhelmed available staff and peer reviewers. Finally, as some researchers moved toward accessing journals electronically through library e-resource subscriptions rather than in print, individual subscriptions to journals began to decline, pinching journals’ profits.

Many stakeholders, academic librarians hardly least, are now saying louder than a whisper that the current situation is unsustainable; access to the scholarly and research record is in decline and cannot be allowed to decline further. A broad-umbrella movement gaining influence and currency is the “open-access” movement, according to which scholarly publication occurs electronically without charging readers for access either directly or indirectly (as through library subscriptions).

In the United States, whose commitment to open access is lagging considerably behind that of other countries (notably the United Kingdom), the federal government has just begun to address access issues surrounding taxpayer-funded research. The National Institutes of Health started a voluntary open-access policy in 2005, requesting but not mandating that funded
researchers deposit some form of their final report in the NIH’s PubMed Central digital repository; the policy’s dismal showing in its first six months of operation, in which only 3.8% of eligible research was actually archived, caused an ongoing re-evaluation of the policy that is likely to lead to mandated deposit of NIH-funded research in PubMed Central. Congress responded also, introducing the CURES Act and the Federal Research Public Access Act (FRPAA), both aimed at increasing open access to federally-funded research.

One form of open access to research bypasses journals altogether, and for that reason is often not included in some definitions of open access that focus entirely on peer-reviewed literature. Frustrated by the slow pace and difficulty of journal publication, some researchers have turned to less formal but more open methods than standard journals of disseminating their results and data. “Grey literature” is burgeoning on the Web: pre-prints, technical reports, conference slide-decks, working-papers series, and more. Although they lack peer review – which has come under attack lately for bias and error in any case – these materials are achieving considerable penetration into the scholarly conversation. Libraries are beginning to ask what role they should adopt in hosting, disseminating, recommending, filtering, and preserving grey literature.

Without peer-review or editorial-board imprimatur, without publisher marketing or librarian support, how are researchers finding grey literature, and why do they read and cite it? Scholars, it turns out, are hardly more immune to the attractions of Web-search engines than undergraduates, and grey literature is present to search-engine crawlers in much higher density than peer-reviewed work.

Indeed, evidence is mounting that research material made available online without barriers to access is more easily found, more often read, and more often and more quickly cited.² This
means that toll-access publishers and researchers are working somewhat at cross-purposes. Researchers desire the widest possible dissemination of their findings; the more they are read and cited and the more demonstrably influential their work, the better for their careers. Publishers whose business model relies on content ownership, on the other hand, must restrict access to the content they own to those who pay for access. Researchers are not generally aware of this disconnect; most have as much access as they believe they need, and when they do not, they tend to blame library budgets rather than publishers. Moreover, the research demonstrating open-access citation advantages has not yet garnered widespread attention among researchers.

In an allied development, “impact factors” that purport (however sloppily or erroneously) to gauge an article or journal’s importance by counting the number of times it has been cited have been gaining influence in retention and tenure hearings. Standard Web metrics such as number of downloads are also catching eyes. Notably, impact factors and download numbers are calculated post-dissemination, shifting quality measurement to some extent from traditional pre-publication measures such as peer-review and acquisitions editing. A piece of grey literature can count its downloads and citations without having to undergo peer review.

How will journals be funded, if not through subscriptions? John Willinsky’s excellent book *The Access Principle* suggests ten economic models for open access (pp. 212–13), countering the myth that the only economic model is that in which the author pays a fee to the publisher to consider or publish the article. Subsidies, grants (either to the journal or to authors), and in-kind support from libraries and other stakeholders may also defray costs. Most intriguing among Willinsky’s suggestions is a publishing “cooperative” (Appendix D, pp. 227–232) in which
libraries, university presses, scholarly associations, and other stakeholders contribute money, expertise, and in-kind support to the publishing enterprise.

Most advocates commonly divide open access to the peer-reviewed (as opposed to grey) literature into two modes, the “green road” and the “gold road.” The “green” road is self-archiving, in which researchers deposit copies of their work on the World Wide Web. Widespread and time-honored in computer science, physics, and engineering thanks to efforts such as arXiv.org, this approach has yet to find numerous adherents in other fields. Self-archiving venues range from individually-hosted author Web pages through discipline-specific repositories (such as library science’s E-LIS [eprints.rclis.org](http://eprints.rclis.org/) and DList [dlist.sir.arizona.edu](http://dlist.sir.arizona.edu/)) to institutional repositories, which are usually managed by the institution’s library. The FRPAA bill is of particular interest because (unlike the CURES Act, which mandates deposit in PubMed Central) institutional repositories may be acceptable deposit venues, depending on how each agency interprets the open-access mandate.

Self-archiving depends on publisher willingness to permit it, or author willingness to flout copyright after signing it over to a publisher. An uneasy and unstable compromise reigns at present. Over eighty percent of publishers require authors to sign over copyright. Most researchers do not understand that publishers own legal rights to their work, although foresighted universities such as MIT are beginning to issue calls for their faculty to amend or cease signing copyright-transfer agreements. Nonetheless, toll-access publishers do not care to alienate their authors by forbidding permission to self-archive or by challenging institutional repositories or e-reserves systems. Some publishers do forbid deposit into disciplinary repositories; curiously, not in fields where disciplinary repositories are well-established.
The “gold” road to open access is publishing journals – with the normal trappings thereof, such as editing and (crucially) peer review – without charging fees to readers or libraries for access. Some of these journals, notably those under the auspices of the Public Library of Science (plos.org/) recoup some or all of their costs by charging authors processing fees, which authors may then charge back to their funders. The sustainability of this model has been questioned, but at this early date it is difficult to distinguish problems specific to open-access from problems common to all new journals, which typically take seven or more years to become solvent.

Of those journals offering open-access content, not all offer open access to the entire journal, in hopes that the embargoed “extra” content will support subscriptions while still offering authors the benefits of open access. Willinsky cites “partial” open access, in which some articles are made available freely while others are reserved for subscribers; other modes of partial open access include reserving editorial material such as book reviews for subscribers, or reserving access to certain file formats to those who pay, as the American Physical Society does with XML-encoded versions of its journal articles. “Hybrid” journals offer authors the choice of traditional toll-access publishing – free to them, not free to readers – or open access supported by an author fee.

Technology [A]

Several technological developments have fostered interest in new publishing models and workflows. Obviously, the World Wide Web reigns supreme; reducing dissemination costs to near-zero is the sine qua non of open access. Other developments should receive due credit also. The steady decline in electronic-storage costs allows cost-effective maintenance of substantial
electronic archives, while the rise in networking speeds permits easy and effective dissemination of multimedia and datasets as well as formal papers. Data curation is now a growth area in digital-libraries research and experimentation.

Software and services supporting lower-cost electronic publishing and archival have seen rapid development since the turn of the millennium. Institutional and disciplinary repositories can choose to install open-source software packages such as DSpace (dspace.org/), Fedora (fedora.info/), and EPrints (www.eprints.org/software/), or buy hosting services such as DiMeMa’s ContentDM (dimema.com/) or BioMed Central’s Open Repository (openrepository.com/). Journal publishers can reduce overhead costs significantly by moving to an all-electronic workflow for manuscript submission, editing, and peer review, using the Public Knowledge Project’s open-source Open Journal Systems (pkp.sfu.ca/ojs/). Cornell University Library is currently testing DPubS (dpubs.org/), journal-management software that is designed to integrate with DSpace or Fedora repository software; DPubS underlies Cornell’s Project Euclid suite of mathematics journals (projecteuclid.org/).

New technical and certification standards for data-modeling, metadata exchange, and archival also underlie the open-access movement. The Open Archives Initiative created the Protocol for Metadata Harvesting (OAI-PMH; www.openarchives.org/OAI/openarchivesprotocol.html), through which data providers such as repositories expose metadata in a straightforward XML format for harvest by Web-crawlers and search engines. OAI-PMH requires that data providers expose Dublin Core metadata, but allows richer metadata of any format (including MARC) as well. The most important OAI-PMH harvesters at present are Google Scholar (scholar.google.com/) and the University of Michigan’s OAIster search...
The National Archives and Records Administration collaborated with the Research Libraries Group (now part of OCLC) in 2005 to publish a draft report, the “RLG-Nara Audit Checklist for Certifying Digital Repositories” (www.rlg.org/en/page.php?Page_ID=20769). The report outlines policy, technical, staffing, and content-management decisions a digital repository must make in order to be considered “trusted.” Although actual certification based on this checklist is some years off, the checklist itself remains useful for repository planners and managers.

One piece of technology that might be thought to impact dissemination and use of electronic materials has proven to be a red herring: the ebook, if by “ebook” is meant a handheld electronic device specifically designed for storing and presenting reading materials. Even during the boomlet of the late 1990s, ebook devices were never designed for or marketed to an academic audience, and current research into “e-ink” has likewise ignored academia. Progress in this realm appears unlikely to impact scholarly publishing within the decade.

Experiments [A]

Experiments in open access have mushroomed both inside and outside libraries since the founding of the Scholarly Publishing and Academic Resources Coalition (SPARC; www.arl.org/sparc/) by the Association for Research Libraries in 1998. SPARC partners with publishers experimenting with more sustainable business models and creates advocacy and educational materials such as the recently-revamped Create Change website (www.createchange.org/).

New open-access journal publishers dot the scene, particularly in fields where journal pricing has become particularly egregious. The United Kingdom’s BioMed Central
BioMed Central (biomedcentral.com), the first open-access journal publisher, manages a large stable of journals in biology and medicine, supporting itself with for-pay services such as image archival.

The Public Library of Science (PLoS, plos.org), which is supported by grants, memberships, and author fees, has launched a small stable of fully open-access journals in biology and medicine. Although a June 2006 hike in author fees caused consternation among some authors and finger-pointing from toll-access publishers, impressively high impact factors for the new journals continue to attract high-quality submissions.

Traditionally toll-access publishers are experimenting with open-access business models as well. Shortly after Springer-Verlag hired former BioMed Central head Jan Velterop as its Director of Open Access, it instituted an “Open Choice” program which charges authors publication fees in return for allowing open access to the resulting article. Elsevier, the Royal Society of the United Kingdom, and other publishers have similar programs.

Several academic libraries now sponsor open-access journals by managing an installation of Open Journal Systems. A few university libraries have gone beyond journals, however: the University of Tennessee has launched Newfound Press (lib.utk.edu/newfoundpress/), which hopes to publish open-access monographs as well as journals. At Rice University, the Digital Library Initiative is one sponsor of Connexions (www.cnx.org), which will host the university’s relaunched university press as an all-digital open-access operation.

Other, generally longer-running, library initiatives offer support for publishing without actually managing the entire endeavor. The University of Michigan Library’s Scholarly Publishing Office (spo.umich.edu), supports online journal and monograph publication through software development, hosting, and conversion and printing services.
Cornell’s Digital Consulting and Production Services (dcaps.library.cornell.edu/) offers a similar service menu, but also includes copyright services such as permissions-seeking and ownership-tracking.

Experimentation with post-publication peer-review is legion. BioMed Central offers the post-publication peer-review service Faculty of 1000 (facultyof1000.com/), in which experts recommend individual papers. Open Journal Systems allows journals to accept commentary on individual articles, archiving the comments along with the article. Naboj (www.naboj.com/) represents a review overlay atop the arXiv disciplinary repository; this model demonstrates how open access enables new services and new communication structures.

The longest-running experiment thus far, however, is the institutional repository. Many academic libraries opened these without fanfare and without solid planning for librarian support, assuming that the gathering momentum behind open access would drive faculty to deposit materials. Reality has proven harsher; faculty are only slowly becoming aware of open access, and have proven unwilling to expend effort to deposit their materials without an external demand (such as a department or university mandate) or an obvious return on their time investment. Libraries have responded with marketing programs and development of such extras as download statistics and “researcher pages” that behave like mini-CVs. Those with the most successful repositories, however, have found library-internal uses for the repository (such as digitized special collections material or electronic theses and dissertations) rather than relying wholly on faculty for materials.

Once we accept that scholarly publishing must change, we have a duty to ask why we should shape its future. Is this in our mission? Are we best-suited to do it? Can we afford to? What’s in it for us?

I believe academic libraries are uniquely situated to redress the systemic imbalances and infelicities in the scholarly-publishing system. Unlike scholarly societies, academic libraries do not expect to fund their other activities via publishing. Unlike for-profit journal publishers, academic libraries have no shareholder obligations; we exist to serve research and researchers. Unlike university presses, academic libraries are not expected to be financially self-supporting. Indeed, open access fits quite neatly into our core mission of disseminating quality information; we should have little trouble justifying publishing expenditures to cost-conscious university administrators.

Moreover, we academic librarians have skills that transfer well to publishing. Our experience with digitization projects has created a fairly large pool of Web-savvy digital librarians. We indubitably understand far more about digital preservation than do publishers. As researchers ourselves and as liaison and reference librarians, we understand the scholarly process and work closely with researchers. We even know all about citation formatting, abstract construction, and indexing!

Such production skills as we lack should not worry us unduly; many necessary links in the publishing chain are filled by the very faculty we are closest to, and most production tasks were outsourced by publishers long ago. Acquisitions editing, editorial-board work, and peer review
are faculty tasks. Copyediting, file conversion, typesetting – all are done by freelancers or service bureaus at commodity prices so low that even volume discounts hardly matter.

We can also offer new publishing-related services that at best would be loss leaders to for-profit or society publishers. Data curation is likely the best lever, particularly in the sciences; integrating it with existing research-data services and institutional repository workflows should draw faculty interest. We can come up with post-publication review services of our own. In the humanities, we can offer production and archival support for innovative digital projects that faculty may not be able to manage on their own. Even the Modern Language Association is beginning to admit that digital scholarship should be valued on its merits just as print scholarship is.5

Scholarly publishing makes or breaks academic careers. Since publication record is crucial to hiring, retention, and tenure decisions, librarians involved in publishing will find that they have real, if subtle, power. Naturally we welcome relevance, but power that differentiates between patron and patron will require some adjustment to our normally patron-neutral outlook. The trust institutions place in their libraries should prove helpful as new publishing projects negotiate the treacherous waters of academic prestige.

Less controversially, involving libraries in the publication process will aid long-term preservation of electronic materials, since the methods by which electronic resources are produced are so closely tied to their suitability for preservation. Leaving preservation of electronic journals to their publishers is an untested and uncertain preservation model. In a toll-access environment, journal backfiles may pay for themselves through library subscription or pay-per-article access models. Deadweight (from a publisher’s point of view) is mounting,
however, from licenses that quite properly insist upon library access to backfiles even after
journal cancellation. When providing access (and associated services such as usage metrics)
costs money instead of making it, how long will publishers remain in the preservation business?
Should they leave it, what will happen to the scholarly record? Libraries must answer this
question for the electronic scholarly record, just as we have answered it for print; experiments
such as the collaborative archival program LOCKSS (www.lockss.org/) and the member-
supported archiving service Portico (www.portico.org/) point to possible futures.

Though considerable attention has been paid to possible subscription savings via gold-road
open access, libraries should factor in other potential cost savings as well. An open-access article
(whether in an open-access journal or self-archived) can be linked to in an electronic-reserve
system without royalty cost or copyright liability. Indeed, this is such a benefit that libraries
might well consider asking e-reserves staff to contact authors of often-used articles to ask that
those articles be self-archived.

The barriers we face in moving toward open access are generally not technological; they are
social and organizational. Librarians are still hashing out libraries’ proper role in judging,
acquiring, and preserving electronic content. Faculty and administrations have been slow to
respond to librarian appeals. Publishers and scholarly societies that own revenue-positive
journals spread fear, uncertainty, and doubt. Finding and preserving the grey literature now
scattered all over the open Web is a tremendous challenge. New open-access journals face the
same marketing, revenue, and prestige issues as any other new journal – except that detractors
tend to attribute difficulties purely to open access. Still, the tide is turning; although a purely
open-access scenario is no more than a chimera, open access makes too much sense not to command a larger piece of the scholarly-communication pie.

**Brave New Open-Access World [A]**

Open access is already changing academic libraries, to judge from the increasing job opportunities for repository librarians and scholarly-communication experts.

Open access is not, however, “free” access, because publishing is not free, even considering unremunerated labor from faculty and the savings in overhead made possible by all-electronic workflows and dissemination. Research funders and institutions are unlikely to accept all the costs of gold-road open access, even in the well-funded fields at the root of the serials crisis. Over time, libraries will have to pick up some costs as well, by joining organizations such as SPARC and PLoS, by offering a pool of funds for faculty to publish in open-access journals, or by offering in-kind services such as journal hosting and archival. At present library budgets are squeezed from both directions, since it is impractical to drop journal subscriptions in favor of support for open-access publications. Eventually, however, open-access experiments both by libraries and publishers should cause subscription price pressures to ease, freeing up money to support open access (as well as languishing monograph budgets).

Many libraries opening institutional repositories have found to their considerable dismay that faculty are not flocking to the new service. Some have responded by allocating staff time specifically to the repository, both for marketing and for services such as batch file import, metadata creation, faculty training, and copyright clearance. Others may decide to close repositories to new deposits while they rethink their planning. In the long term, however, I expect many libraries to decide that not every institution needs its own repository. Consortial
repositories such as the Washington Research Library Consortium’s Aladin-RC (aladinrc.wrlc.org) conserve scarce technical staff, allowing each individual library to focus on outreach and services.

To some extent, whether faculty embrace open access is out of libraries’ hands, dependent as it is on decisions by government, private research funders, scholarly organizations, and college and university administrations. Some faculty will have no choice, owing to mandates from research funders or institutions; we should stand ready to explain what is happening and assist faculty as they start to comply with these mandates. Other faculty will find open access on their own, enticed by hopes of additional professional impact or librarian curation for their work; we should encourage, support, and recognize them for their efforts. Skeptics are legion and will remain so; we must express our positions firmly but respectfully, and be transparent about our role in the open access movement and what we hope to gain from it.

Certainly libraries, library consortia, and library professional organizations should try to influence decisions by funders, institutions, and governments in favor of open access. Less-well-funded libraries have a particular interest in speaking up, as their access to research literature will improve most, and their experience counters the unfortunate refrain from short-sighted large research institutions that open access is unnecessary because their faculty have all the access they need. Immediate results are unlikely, however; wide-ranging changes in a long-standing publication system take considerable time. Self-archiving mandates are probably out of reach of most institutions at this juncture. Asking (or even requiring) faculty to retain some rights in their journal articles, a strategy pursued at such universities as MIT and the University of California system, has met with more success, and lays important groundwork for further progress.
Libraries can expect software for journal management and archival to improve significantly. DSpace, for example, is less than five years old. Repository and journal-management software, with its emphasis on preservation and workflow, is a natural complement to digital-library software such as Greenstone (greenstone.org/), which emphasizes attractive and functional user interfaces. I expect these different types of software to interoperate better in the short term; strawman proposals are already being floated for new interoperability layers. In the long term, they may merge, become modular enough to give rise to mix-and-match systems tailored to specific needs, or adopt so many features from one another as to be nearly indistinguishable. Either way, we can expect more finished, flexible, reliable, and capable software than we currently have.

We can also expect further experimentation with shared data and metadata. The rather crude search and discovery systems OAI-PMH has made possible point the way toward much more refined systems that mingle advanced text-mining techniques with the collection developer’s keen eye for value and the cataloguer’s and indexer’s long-honed subject-assignment skill. Easy availability of datasets will likewise jumpstart innovative and necessary meta-research in the sciences.

Not all harvesting experiments with OAI-PMH have been rosy, of course. The National Science Digital Library discovered that considerable technical support was necessary for data providers to reach even a minimal level of technical efficacy and metadata adequacy. Clearly, academic libraries will have to bear metadata-training and metadata-revision burdens; we cannot possibly expect to keep up with the volume of metadata creation all by ourselves. Together with Indiana University, the University of Illinois at Urbana-Champaign secured an IMLS grant to do
low-cost training across the country in 2007 and 2008 on creating “shareable metadata.” Their experiences will doubtless inform further training programs.

The greatest advantage of open access to my mind is that it will inevitably awaken researchers to the economic consequences of their own publishing decisions. The current model separates researchers’ decisions about where to publish from libraries’ decisions about what to buy. Arguably, this exacerbated the serials crisis, as researchers with eyes on their own career advancement ignored or blamed librarians for overburdened library budgets. My hope is that the next stage of open-access growth and development will intertwine libraries further with the research process as publication partners, co-creators and disseminators of information as well as its collectors, describers, and preservers.

Bio [A]

Dorothea Salo is the Digital Repository Services Librarian at George Mason University. She is responsible for technology development, maintenance, policy, metadata, and outreach for the Mason Archival Repository Service (mars.gmu.edu/). She holds a master's degree in Library and Information Studies and another in Spanish from the University of Wisconsin at Madison, and a BA in comparative literature and Spanish from Indiana University at Bloomington. Her previous experience includes typesetting and SGML production for scholarly books and journals, and work on technology standards for electronic books.


6 See the new MIT Copyright Amendment Form, libraries.mit.edu/about/scholarly/copyright-form.html (accessed July 6, 2006).

7 See the University of California Academic Senate’s white papers and policy proposal at http://www.universityofcalifornia.edu/senate/committees/scsc/reports.html, accessed July 6, 2006.
