

From the University Presses — How to Establish a Research Agenda for Scholarly Communication, Part II: A Sympathetic View

Column Editor: **Sanford G. Thatcher** (Director, Penn State Press, USB 1, Suite C, 820 N. University Drive, University Park, PA 16802-1003; Phone: 814-867-2220; Fax: 814-863-1408) <sgt3@psu.edu> www.psupress.org

As promised in Part I of this essay, titled “The Paranoid View,” Part II will look at the *ACRL Report* (November 5, 2007) through a set of different lenses. In Part I, I questioned the report’s “characterization of the need for policy reform in ‘legal matters’ and its resetting of priorities toward a greater role for libraries in the research processes and dissemination of their results as well as its support for modes of informal communication over formal publication.” Here, while remaining wary of what the *Report’s* authors view as “balance” in copyright law, I want to offer a less jaundiced assessment of the “resetting of priorities.”

Conversations with several librarians at **Penn State** and elsewhere, supplemented by reading of some recent and forthcoming publications, have given me a better appreciation of the reasons why librarians feel some obligation to tackle the daunting job of supporting and preserving the records of scholarship throughout its entire life cycle including, significantly, both data sets of a wide variety of types and informal as well as formal modes of communication. It is not just that technology has made it possible to track and retain much that was regarded hitherto as outside the scope of libraries’ roles as collectors and organizers of the products of scholarship; it is also because there is value perceived by scholars themselves in having long-term access to such materials relating to the process of scholarship, not just to its final state in published form. And the new avenues scholars are able to pursue in their investigations are even beginning to blur the lines between primary and secondary sources and between informal and formal modes of communication.

With respect to data sets, **Christine Borgman** presents the emerging picture well in her recent book *Scholarship in the Digital Age: Information, Infrastructure, and the Internet* (MIT, 2007): “The data deluge is affecting scholarship and learning in ways both subtle and profound. Producing great volumes of data is expensive, whether by scientific instruments or from national or international surveys. Larger teams of researchers are collaborating to produce these data sets. More funding agencies, journals, and conferences expect researchers to make their data available for others to mine. Sharing data is seen as a way to leverage investments in research, verify research findings, and accelerate the pace of research and development. In some fields, the data are coming to be viewed as an essential end product of research, comparable in value to journal articles or conference papers. Another trend is the blurring of the distinction between primary sources, generally viewed as unprocessed or unanalyzed data, and secondary

sources that set data in context, such as papers, articles, and books. Data sets are being listed as scholarly publications in academic vitae and cited as bibliographic references in scholarly articles. Scholarly publications may contain embedded data sets, models, moving images, and sound files, and links to other documents, data sets, and objects. Systems to manage scholarly documents must accommodate much more than text, tables, and figures” (pp. 8-9).

Librarians have been well ahead of publishers in thinking about how to deal with data sets as part of the scholarly record, but the day is not far off when publishers also will be drawn into this discussion. One advisor to university presses, **Joseph Esposito**, has urged them to try increasing their role in scientific publishing not only by taking advantage of their presence on campuses to detect new and emerging areas of science before commercial publishers find out about them but also by focusing more on the primary materials of science like data sets and less on the secondary analyses built on them. And just at the end of 2007 an invitation was extended to the libraries, presses, provosts, and IT divisions of the universities of **Illinois**, **Michigan**, and **Penn State** to collaborate in a new effort to support “data communities” in the humanities and social sciences. A background statement for this project reads in part: “A decade ago linking references (CrossRef) revolutionized scholarship. Today, the scholarly community faces the grand challenge of building robust and resilient digital data frameworks for preservation and access to the resources and products of the digital age. Linking data and documents will be among the great benefits of a distributed infrastructure for scholarship, and although the initial national focus concentrates on science and engineering disciplines, addressing the emerging needs of the social science and humanities disciplines also presents great challenges and great opportunities to enhance and advance scholarship in these broad disciplines. The shift to digital data both enables and mandates a more active role for the domain scholars in the data publication and curation process. Without their active involvement it will be too onerous to curate large amounts of data, and information that is important to the community but not to the individual researcher will be lost. We require a socio-technical infrastructure that will encourage scholars to properly annotate primary and secondary data they create and to capture data that are now discarded byproducts

of their research. **NSF’s DataNet** program will fund the foundational steps to building a sustainable digital data preservation and access network. By focusing on the humanities and social sciences and on the data relating to the publications formally produced by our universities, this project will be at a scale more easily tackled by our library and university press staffs along with subject domain experts and information technologists and will result in an exemplar curated collection of digital content while building a community of users.”

But, even within traditional published scholarship, changes are happening that require a readjustment of our attitudes toward what counts as informal and formal scholarship. **Karla Hahn** pointed me to a recent **National**

“Librarians have been well ahead of publishers in thinking about how to deal with data sets as part of the scholarly record, but the day is not far off when publishers also will be drawn into this discussion.”

Bureau of Economic Research paper by **Glenn Ellison** titled “Is Peer Review in Decline?” (No. 13272, July 2007) that reveals an interesting trend: “Over the past decade there has been a decline in the fraction of papers in top economics journals written by economists from the highest-ranked economics departments. This paper documents this fact and uses additional data on publications and citations to assess various potential explanations. Several observations are consistent with the hypothesis that the Internet improves the ability of high-profile authors to disseminate their research without going through the traditional peer-review process” (Abstract). Surely, papers by such distinguished scholars published thus informally can hardly be ignored just because they do not appear in branded journals. And at a recent discussion with librarians at **Ohio State**, where **Nancy Eaton** and I were invited to talk about the future of scholarly publishing, one professor of law noted that there are now some senior scholars in his field who are attracting considerable attention to their blogs, which are regarded as significant contributions to ongoing scholarly debates. Here, too, a seemingly informal mode of communication is taking on aspects of formal scholarship. Librarians have good reason to want to capture such strands of scholarly discussion and make them as permanently available as traditional monographs and journals.

continued on page 65

But, in assessing what is worth capturing and preserving for the long term, librarians are taking on a role that is less familiar to them than it is to editors at university presses whose primary job it is to help sort out the wheat from the chaff of scholarship and brand the best as worth the investment of funds to make it permanent in the form of publications. Some librarians are already writing about the virtues of combining the strengths of libraries and presses synergistically, building on a prominent theme from the **Ithaka Report**. My **Penn State** librarian colleague **Michael Furlough**, writing about “University Presses and Scholarly Communication: Potential for Collaboration” in *College & Research Library News* (January 2008), offers the following helpful analysis: “Libraries should care about the health of university presses because publishers and publishing-related services are crucial to libraries’ own future. Many librarians now help students and faculty use digital content and technologies in their research and teaching, and we are supporting them in elaborating new and transformative uses of these materials. Increasingly we support more parts of the entire process of scholarship, and, especially in newer media, we are expanding our services to the process of authoring and creation, and then linking that to the process of presentation and archiving. Libraries have invested significantly in technology platforms to manage, provide access to, and (in time) preserve large digital collections. But presentation means dissemination, not publishing of research, and librarians need to understand the scope of both to support scholarly communication more effectively. Our principles of selection — for the materials we buy or license — are based on service to our local faculty and students, not on the same editorial principles that guide publishers. We think of our clients as ‘users’ or ‘customers’ rather than as ‘producers’ and ‘authors,’ but the latter identities are more important to them in establishing their career path. Our attempts to collect their research in institutional repositories could perhaps be more successful if we think of their needs as scholars and producers of research, not just users of our reference and archiving services. Publishers and university presses may know little about how our faculty conduct research, but they know much better than we do how to cultivate their scholarship and bring it to light” (p. 33). **Furlough** goes on to note: “Both libraries and university presses are losing a large part of the authority they have held as arbiters of quality and channels for content access as those roles have migrated to other agents. The real opportunity in collaborations between presses and libraries lies in sharing risk and leveraging their wagers on the future of scholarship in the academy. By linking up the processes of scholarly creation with access and stewardship, libraries have an opportunity to truly attend to the entire life cycle of scholarship. The primary materials in our archives are the future datasets for humanists and social scientists, and our publishing colleagues can help us analyze our

markets, think through our own principles of content selection, and identify opportunities for added value, especially when it comes to identifying and selecting the stuff that **Google** isn’t planning to scan. It’s easy to talk about what scholarship of the future might look like: dynamic, networked, immediately accessible, and quality-controlled through computational systems as well as human assessment. But we don’t know all the small steps to get there, and we need more partners to help us do so — and not all of these partners should be found in our computing departments and IT organizations. Both of us [libraries and presses] are redefining ourselves, and we both need to refocus on all the core elements of scholarly creation and communication to understand the whole cycle more completely. We can’t do that independently in libraries, and university presses bring value and needed expertise to our profession’s attempt to assert new roles in relation to publishing” (pp. 34-35).

This is a spirit of collaboration I can fully endorse, and to the extent that this was the aim of the **ACRL Report**, I applaud its goals and hope its invitation to continue the dialogue will be accepted by members of the university press community. If I have any lingering worries, they arise from the **Report’s** recommendation (p. 14) to “study the potential cost savings of reducing the acquisition, processing and shelving of print books and journals to reallocate funding to digital content creation and preservation” in conjunction with this powerful reminder from **Clifford Lynch**, an advisor to the **Report’s** authors, who wrote recently in an article titled “A Matter of Mission: Information Technology and the Future of Higher Education” (in Richard N. Katz, ed., *The Tower and the Cloud*, 2008): “In the print era, primary stewardship of the record of scholarship was very closely tied operationally and economically to the dissemination system (publishing);

research libraries purchased this record, made it accessible, and preserved it. The system of research libraries, and the broader system of organizations that managed the base of evidence for scholarly work, represented a substantial and sustained investment both by higher education and by society as a whole. The growth of new kinds of scholarly communication today, the move to e-research, the reliance of scholarly work on a tremendous proliferation of data sets (some of them enormous) and of accompanying software systems threaten to greatly increase the cost and complexity of the stewardship process and to at least partially decouple it from (traditional) publishing, meaning that libraries need to reexamine and redefine their roles appropriately to address these new scholarly works and this new body of evidence for scholarship. Commitment to activities like data curation and management of faculty collections will increasingly characterize research libraries as much as the comprehensive collecting and preservation policies for published literature and personal papers. The cost of stewardship is, I believe, going to rise substantially.”

With the rise in cost will surely come even tougher decisions about how to allocate scarce resources. The recommendation of the **ACRL Report** to consider diverting funds away from print to digital collections cannot help but increase the insecurity of university presses, which have largely succeeded in transitioning from print to digital in journal publishing but have yet to figure out a way to do it successfully for monographs, though experiments are under way. 🌱