

The following paper was delivered at a program entitled, "Scholarly Journals: Our Futures in the Digital Soup," presented by the Council of Editors of Learned Journals on December 30, 2004 in conjunction with the Modern Language Association Annual Convention.

The Future of Scholarly Communication in the Humanities: Adaptation or Transformation?

Richard K. Johnson
Director, SPARC (Scholarly Publishing & Academic Resources Coalition)

All around us enterprises are weighing the demands of adapting or transforming their business models against the risk of ultimate extinction. Music, motion pictures, retailing, travel, telecommunications – giant changes are underway in these industries and many others. For agile players, evolving consumer demands may present new opportunities. Others will be left behind like buggy whip makers as new entrants emerge.

But we're interested in scholarship, not commerce. Are the rules different here? Perhaps. But the forces driving change are the same. Technology will remake scholarship no less than industry. Publishing will adapt to or be transformed by the possibilities of a ubiquitous network. So will libraries.

Scientific publishers were quick to recognize that the low marginal cost of dissemination on the Internet could drive gains in revenue and profit. My own experience of the past half dozen years with SPARC has been mainly in encouraging the potential of the network to serve the academy by maximizing the extent of access to research while reducing its cost. My remarks today are colored by my experience in scientific rather than humanities publishing.

Of course, the humanities and science are quite different. Journals are the primary means of publishing in the natural and physical sciences, while monographs play a more central role in the humanities. Moreover, digital technologies and the network have been taken up at a slower pace in humanities than in science, in part due to cultural barriers. However, this also might be due to the reduced availability of investment capital in the humanities. In most areas of science, investigation is inherently more expensive than in the humanities. Enhancements to infrastructure that make electronic dissemination feasible in scientific fields have often been subsidized through federal government spending on scientific research.

Take the National Institutes of Health, for example. NIH is a \$28 billion federal agency, and the vast majority of its budget supports extramural research grants. Recently, with active encouragement from Congress, NIH proposed a groundbreaking plan to make articles on NIH-funded research available free to the public. As many as 60,000 peer-reviewed articles could be added to NIH's online digital archive each year. The projected cost is between two and four million dollars per year — a pittance relative to the size of NIH's investment in research. The cost is so low in part because the required infrastructure already exists at the National Library of Medicine's National Center for Biotechnology Information.

The most compelling motivation for Congress and NIH to support open public access to NIH-funded research is to demonstrate to taxpayers the return on their investment in doubling the size of NIH in recent years. Perhaps there is a lesson here for the humanities. Expanded public exposure for scholarship in the humanities offers the potential for enhanced political and financial support. By reaching a broader audience beyond specialists in a single field, open global sharing of knowledge also will support interdisciplinary scholarly perspectives.

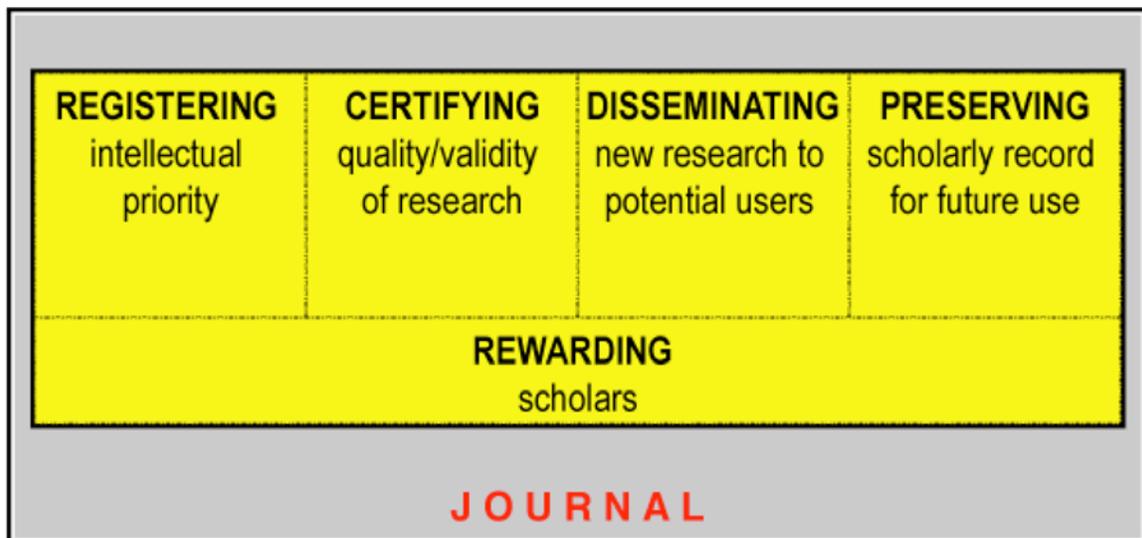
One thing that science, the social sciences, and the humanities have in common is the growing problem posed by the exploding volume and price of information. Libraries cannot afford access to the full range of information needed by researchers. Budget strains are forcing them to forgo the purchase of new journal titles, cancel subscriptions altogether, and reduce the purchase of monographs. The resulting crisis touches not just libraries and readers but also the publishing programs that rely on library spending and the authors who depend on publishing for the recognition that drives their careers.

A growing chorus of observers has concluded that the current system of scholarly and scientific publishing is unsustainable. This situation — coupled with the potential of the Internet — makes change inevitable. Frankly, I don't know exactly how this change will play out in the humanities or how closely the experience will mirror the sciences. But to understand the potential for change, I suggest we examine the fundamental purposes of scholarly publishing.

It has been suggested that scholarly publication performs several enduring functions:

- *Registering* the intellectual priority of an idea, concept, or research;
- *Certifying* the quality of research and the validity of a claimed finding;
- *Disseminating* new research to potential users;
- *Preserving* the scholarly record for future use; and
- *Rewarding* scholars for their work.

Figure 1. Traditional printed journal



The traditional print publishing model (Figure 1) integrates these functions in journals. Each of the functions actually comprises multiple elements of the scholarly publishing “value chain” – the many individual processes that cumulatively add value to information as it moves from creation to consumption. These publishing elements—including production, editorial processing, and distribution—have a significant impact on the costs of scholarly journals under the existing system. The integration of these many processes in journals has several implications.

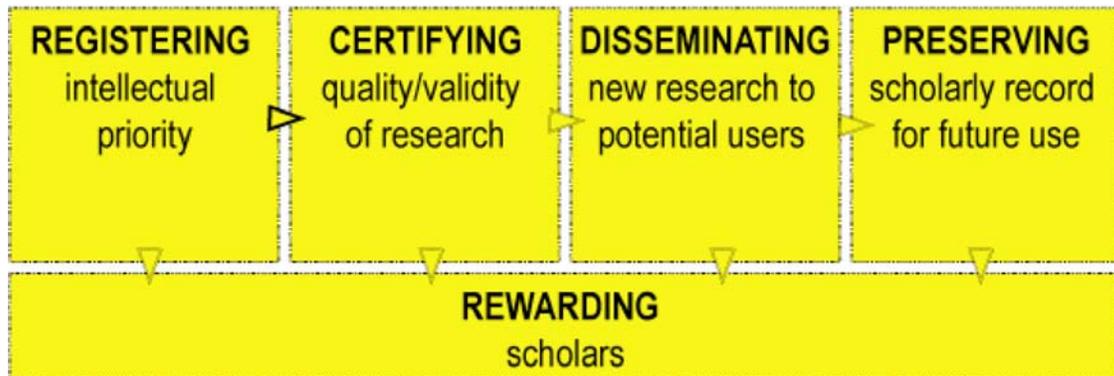
First, the melding tends to disguise the fact that most of the direct labor and much of the indirect cost required for each of the components derives from the academic institutions that ultimately bear the cost of scholarly journal subscription fees. Consequently, the value of the journal may derive less from the cost of creation than from its monopoly control over unique articles — conveyed via copyright assignment — and from the prestige associated with the research it certifies.

A second implication of this integrated publishing model is that increased cost efficiencies in any one component of the scholarly publishing value chain do not translate into reduced journal prices. But deconstructing the value chain, fragmenting it into multiple businesses, makes the value added by publishers more apparent and separable and helps eliminate inefficiencies.

For example, a key publisher contribution to the scholarly publishing value chain historically was the distribution component: typesetting, printing, marketing, and fulfillment were specialized and expensive tasks that authors and libraries gladly delegated to publishers. With the evolution of digital publishing and networked distribution technologies, the relative value of print production and distribution has declined. Not surprisingly, most publishers are unwilling to accept the commensurate decline in revenues and profits that their reduced participation in the chain would yield. Therefore, many have responded with real or artificial added-value programs, such as bundled print-and-digital offerings or cross-subject aggregations, to support prices.

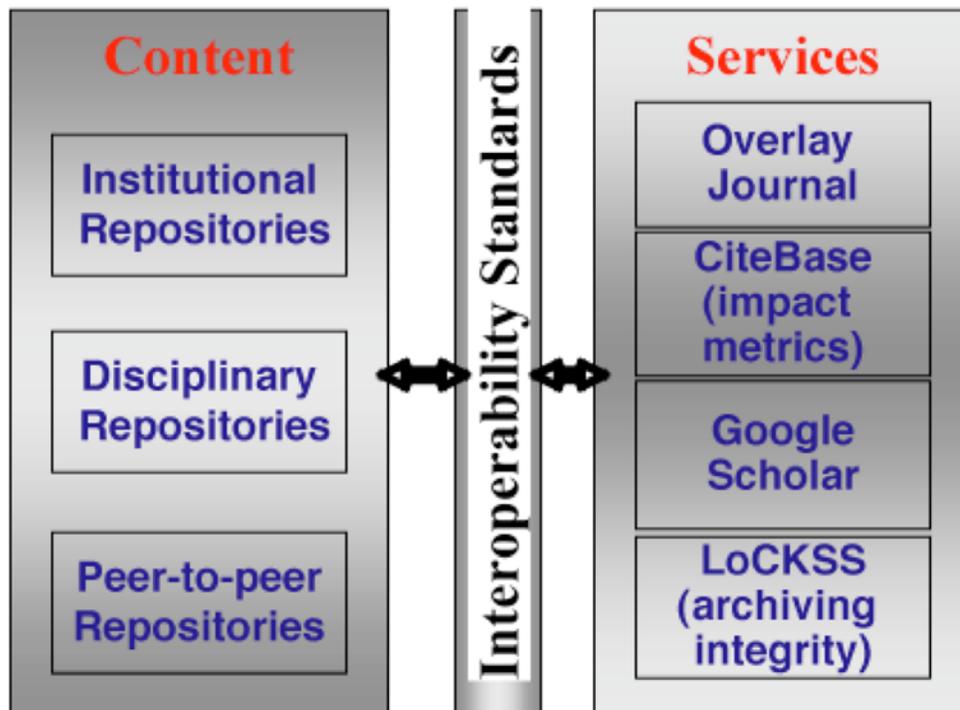
Digital publishing technologies, extensive global networking, an increasing volume of scholarship, decreasing satisfaction with a dysfunctional economic model — all these factors are changing the fundamental structure of scholarly publishing. They allow and provide impetus for its components to be de-linked, both functionally and economically (Figure 2). When the functions are unbundled and begin to operate separately, each is more efficient and competitive. The bundled structure of traditional journals can be replaced by a system of cooperating distributed agents, responsible for different aspects of registration, certification, dissemination, preservation, and rewarding.

Figure 2. Deconstructed journal



Fundamental to implementing this unbundled model — the deconstructed journal, if you will — is the logical separation of the content and service components (Figure 3). This separation allows for distributed open-access content repositories to be maintained independently of value-added services fulfilled discretely by multiple service providers. Once the components of scholarly publishing are logically separated, the registration, certification, dissemination, preservation, and rewarding functions can be undertaken by any organization with sufficient resources, intellectual prestige, organizational standing, and market position.

Figure 3. Content-Services Relationship



This is all rather abstract, but we can see it beginning to play out in the real world. I mentioned earlier that the National Institutes of Health plans to offer access to articles based on NIH-funded research in their online digital repository, PubMed Central. NIH grantees will be asked to provide the agency with electronic copies of their manuscripts upon acceptance for publication – that is, the author’s final word processing file after all modifications due to the peer review process. NIH will delay offering public access to these files for six months as a means of protecting journals from cancellation. In biomedicine, six months is an eternity, so this ensures that subscribers won’t substitute free PubMed Central access for their paid journal subscriptions. Moreover, the proportion of articles in a typical journal that are *not* based on NIH-funded research — and hence will not be in PubMed Central — is sufficient to further protect subscriptions.

This plan will have the effect of disaggregating the dissemination and preservation functions. PubMed Central will serve as a supplementary dissemination channel and will preserve NIH’s research over the long-term. It will not certify research quality, except by virtue of including only peer-reviewed research.

Something similar has existed for years in physics with the now-famed arXiv.org digital e-print repository developed at Los Alamos National Lab and now residing at Cornell University. ArXiv.org — which today also serves mathematics, non-linear science, computer science, and quantitative biology — coexists with the corresponding journal publishing industry. Unlike NIH’s PubMed Central, arXiv.org accepts pre-prints – articles that have not yet been peer-reviewed. In biomedicine this would be considered a threat to public health, but there are no such qualms in physics, math, and the other fields served by arXiv.org. There, open sharing of preliminary texts is used to obtain collegial input on an article prior to its submission for publication. Even though authors often replace their preprint with the final version of the article, arXiv.org and traditional journals have sufficiently differentiated themselves so that neither threatens the existence of the other.

Another sign of disaggregation is that a growing number of universities are developing their own open digital archives as a means to harness the intellectual capital of the institution. Commonly known as institutional repositories, they build on a growing grassroots faculty practice of self-posting research online. Institutional repositories necessitate that libraries—as the institution’s logical administrative proponent — facilitate development of university intellectual property policies, encourage faculty authors to retain the right to self-archive, and broaden both faculty and administration perspectives on these issues. Nevertheless, they can be implemented without radically altering the status quo.

PubMed Central, arXiv.org, institutional repositories — these all complement the existing scholarly publishing model. But alone they do not replace it because they do not take on the critical certification or rewarding functions. Until there are credible, disaggregated means of certification and rewarding, traditional journals will endure. We are seeing some experimentation now with what is often termed “overlay journals.” These are services that perform peer review of selected preprints and link readers to the texts in a digital repository. They look like any other online journal – some even have print components – but they explicitly acknowledge the disaggregated environment in which certification, rewarding, and dissemination can occur independently.

There is burgeoning interest today in open-access journals, which are a bit like open repositories with peer review. Although the *Directory of Open Access Journals* lists more than 1400 peer-reviewed journals that are free to readers, probably the best known is the

Public Library of Science's PLoS Biology. This functions much like a traditional high-prestige journal — except that it does not charge users for access, it does not require that authors transfer their copyright to PLoS, and it gladly deposits the final published texts of all articles in NIH's PubMed Central for preservation.

PLoS's model might be thought of as disaggregated because the publisher has chosen to apply the Creative Commons Attribution License to all works it publishes. This broad license was developed by Stanford Law professor Lawrence Lessig to facilitate open access to, and free use of, original works of all types. While authors retain copyright to their work, they grant users the right to copy or distribute it or to make derivative works provided the original author is credited. This means that the dissemination function can take place outside the context of PLoS's journal.

What do these examples have to do with the humanities? They simply suggest that powerful — if not unstoppable — forces are chipping away at the traditional journal. Will the outcome be adaptation or transformation? I think there will be both. There will be e-journals (and e-books) that look much like what is supplied in print today. But the foundation beneath these, the ways in which they are accessed and used, what they contain, and the profile of users is likely to be transformed.

The toughest issues we face today revolve around business models – who pays the tab in a disaggregated environment? Perhaps toughest of all, how is the certification process supported?

Publishers and libraries aren't the only players asking themselves these questions. The costs associated with publishing are the least part of the overall research process. Since academic institutions, funders, and the public are key beneficiaries of research, I think we can expect them to play new, active roles in reshaping scholarly communication in the sciences, the social sciences, and, yes, the humanities.