

**CRISES AND OPPORTUNITIES: A SCIENTIST'S VIEW OF SCHOLARLY  
COMMUNICATION**  
**A White Paper for the UNC-Chapel Hill Scholarly Communications Convocation**  
**January 2005**

by  
Robert K. Peet, Department of Biology & Curriculum in Ecology

**1. The role of referred journal articles in the sciences**

In a 1998 review of scholarly publication in the digital age, Vince Resh\*\* wrote, "Research articles in refereed journals are the traditional "coin of the realm" for academic scientists. Through their publications scientists either become known or remain unknown. Moreover, their initial appointment and eventual tenure, promotions, and research funding are largely based on the quality and the quantity of their publications." Seven years after his review traditional journal articles retain their central role, but we can see changes on the horizon driven by changing technology and disciplinary culture.

Before looking into the evolving role of journal articles, it is important to step back and review their functions in scholarly communication and the degree to which those functions are tied to the traditional journal article. Karen Hunter of Reed-Elsevier once described four traditional functions, roughly as follows: 1) certification, which establishes that the author had the ideas expressed and establishes a dated claim for priority; 2) validation, where the peer-review system verifies the quality and correctness of the content; 3) awareness, where the journal serves to disseminate and advertise the content; and 4) archiving, which is the guarantee of long-term access. All of these functions are critical to scholarly communication and must be included in any new models we might adopt. Preprint archives (see section 8) are now arguably the major mode of communication in physics and mathematics and are highly efficient and preferable to journals for certification and awareness, but can fall short in validation and archiving. Conference proceedings are now a highly visible alternative to journals in some areas of engineering and technology, though perhaps weaker in awareness and archiving. Institutional repositories are widely touted as a solution to the "library crisis", but university repositories will have difficulty achieving the reputation for peer review that an international, field-specific journal can achieve.

**2. Digital publication – the new standard**

During the period 1993-1999 the UNC-CH Couch Biology Library recorded an average of approximately 500 uses per year of the journals of the Ecological Society of American, where a use is a volume or issue taken off the shelves and either checked out or left for reshelving. During the period 1999-2002 the number of uses dropped to an annual average of around 150, reflecting the availability of older issues in the digital JSTOR archive. We now have full digital access to all issues. During the academic year 2003-2004 there was a total of 7 uses. Whereas five years ago many predicted a gradual change to digital access, the shift took place almost overnight. The dramatic transition to digital access, together with new online searching tools has had a parallel impact on library usage. Over the past fourteen years there has been a drop of

approximately 50% in annual circulation and 70% in reference questions in the Couch Biology Library.

Today nearly all young scientists read journals almost exclusively in digital format, and are even reluctant to look for articles from “primitive” journals that are not yet available in digital format. Journals that are slow in making the transition are precipitously losing market share as measured by citation statistics. It is no longer economically viable to publish a strictly paper journal. Further, we should not expect this wave of transition to stop with scientific journals. The writing is on the wall – we are headed toward completely digital dissemination of scholarly communications. Already there are major projects underway to digitize collections of books and journal backruns (e.g., [JSTOR](#), Stanford & Google<sup>\*\*</sup>). Science is a bit ahead, but we should expect this transition to spread quickly to other fields as well.

Only a few years ago concern was widely expressed that digital journals were not as strong and rigorously peer reviewed as print journals, and might not count for as much in tenure and promotion decision. What a difference a few years makes for our perspective. Paper copies of journals are being moved off campuses and may disappear altogether. Journal prestige will continue to be critically important, but biases against strictly digital formats are rapidly fading.

There are possible secondary implications of the transition to primarily digital access. Production costs will drop, even if we cannot count on the savings being passed on (production of the paper copy generally runs around 35% of the cost in a traditional journal with significant staff investment in editorial services, and can be much higher in small journals run largely on volunteer labor). The combination of digital distribution and digital processing of manuscripts should bring a significant drop in the time lag between submission and print, and where it does not journals will face mounting competition from preprint servers as a mechanism for rapid dissemination. Finally, there are widely discussed but largely unsettled concerns over long-term preservation and archiving of digital material that must be resolved.

### **3. Layered publications – digital publication opens up new opportunities**

The transition to digital communication has brought with it opportunities for new formats and content. Numerous changes have already started to appear. Digital journals now routinely have references hyperlinked to their sources via embedded digital object indicators so that it is possible to bounce back and forth through a web of interconnected articles. Many journals have links embedded for other features, such as scientific names of organisms being linked to the taxonomic treatment in [ITIS](#), and DNA sequence data being linked through [GenBank](#) accession numbers.

Numerous journals have created linked archives for supplemental material available exclusively in digital format. For these journals, appendices typically no longer occur in the print version of the journal, but are confined to the archive. In addition, supporting datasets and other material can be filed in the archives. The opportunity presented by digital archives to publish important but rarely consulted supporting material represents a major advance and will allow reanalysis of data using novel methods, as well as the compilation of primary data from multiple papers for heretofore impossible meta-analyses.

Digital archives and papers rich in hyperlinks represent only the first step in a transition to what I think of as layered articles. As journals become exclusively digital we can expect inclusion of linked content that is too expensive or impossible for paper format such as extensive use of color figures and photos, sound recordings, video clips and animations, datasets, computer code, attached comments contributed by readers, and many others. I anticipate the first layer of future articles to be short and to the point, conveying the essential findings and their implications. These would be essentially expanded abstracts or digests. Details of methods, analyses, and literature will be on a lower level accessible via hyperlinks. The future paper journals will likely include only this first layer and serve a current awareness function, much like a newspaper. Some important journals like *Science* are already heading in this direction.

#### **4. How we find information**

When I was a graduate student, we kept abreast of the literature by reading new issues of journals as they came out, and we found supporting literature through the reference sections of the papers. We often subscribed to the most important journals in our field so that we would get them quickly. Today students and young professionals use tools like Science Citation Index to find any important article on a topic via a keyword search. They search forward through articles that have cited critical foundational papers, and they search for the most closely related papers in terms of citation overlap. They subscribe to services that send them announcements of new articles that cite papers of interest to them, and they create their own personalized notification services based on key words and papers of interest to them. This has brought a dramatic change in culture where students are no longer “loyal” to a few journals or a discipline, but instead end up reading only one or a few articles from a journal issue, but read articles across many journals and even many disciplines.

The methods of discovery I describe above are extremely powerful and allow much more efficient discovery than was possible in the past. However, the essential tools are often expensive. UNC-CH pays [ISI](#) an annual fee of over \$100,000 for access to Web of Science plus over \$50,000 more for access to [BIOSIS](#) and [Journal Citation Reports](#). As a consequence of the high prices, these types of tools are often unavailable at smaller campuses leading researchers to employ lower-cost alternatives. A particularly interesting alternative that appears to be rapidly gaining usage and which provides at least some citation information and linking is [Google Scholar](#), available free to everyone.

#### **5. Data archives, registries and discovery tools**

Articles and other traditional print publications do not constitute the only major type of information that scientists would like to communicate or discover. More and more, we are searching cyberspace for data in support of our research questions. This is where the information landscape is changing most rapidly. Archives for data and other information, as well as tools for discovering those data and merging them for new analyses, are just beginning to appear.

NSF tells us we have a responsibility to make our data available digitally to other workers, but they have not yet told us how to do this. What they have done is fund a significant number of projects to develop the infrastructure to archive, locate, combine, and analyze data collected over large and dispersed information grids. In some cases standards and mechanisms existed for sharing data, such as for gene sequence data and museum specimens of organisms, but these are

the exceptions. In many cases the standards for documenting and archiving data are just being developed. Examples of discipline-specific cyber-infrastructure projects include [SEEK](#) which takes on the challenge of modeling, designing, and implementing data discovery, integration and visualization components for a semantic web in ecological and environmental science, and [GEON](#) which aspires to do the same for geosciences.

One key component of the future world of data sharing and grid computing is certain to be dataset registration where datasets conformant with standard metadata mark-up requirements are registered so that they can be efficiently found, searched, and mined across the web. These datasets might reside in archives maintained by journals, professional societies, government agencies, or in institutional repositories (for this function there is real potential for institutional repositories, more so than as homes for in-house articles). A number of initiatives are underway where professional scientific societies are collaborating to develop data sharing standards and data registries. Of course, it takes no small amount of time and effort to mark up raw data in a form conformant with emerging standards. The primary motivation will come in the form of requirements for data archiving on the part of funding agencies and journals. A secondary motivation for some will be the new opportunities for collaboration, data preservation, and dataset citation.

#### **6. The library crisis – the current trends are not sustainable**

For nearly 20 years I have been participating in conferences on “The Library Crisis”. Despite the urgency and desperateness of the situation libraries face, it is difficult to continue to justify calling the same phenomenon a crisis for two decade; perhaps it would be more accurate to view the “crisis” as an unsustainable economic system. The problem is largely a consequence of commercial publishers aggressively establishing hundreds of new journals and steadily ratcheting up their prices. Prices of journals published by commercial firms have been increasing at a rate roughly three times the inflation rate. Consequently, libraries typically pay 4-6 times as much per page for journals owned by commercial publishers as for those published by professional societies.

Despite the high prices of the commercial journals, they are often of lower quality or importance than those published by professional societies. Bergstrom & Bergstrom ([2002](#)) conducted an economic analysis of ecological journals (and several other fields with similar results) wherein they observed that in the year 2000 a librarian could purchase subscriptions to all of the ecology journals listed by [ISI](#) for \$55,000, but could purchase half of the pages for a mere \$12,000. They went on to observe that if a librarian were trying to optimize cited articles, she could purchase journals responsible for 50% of the citations for under \$5000.

The most egregious price increases have taken place in the sciences, so at least for the present the burden of finding a solution largely falls on the scientific community. Clearly doing nothing is not an option. If the present trend continues, we will fail in our basic goals of creation and dissemination of knowledge: creation because we will not have access to critical resources; and, dissemination because no one will be able to afford to read our results. Possible solutions to the library crisis are relatively limited. I have repeatedly heard reference to five options: 1) behavior modification, 2) open access, 3) decoupling dissemination from validation, 4) retention of copyright and license rights, and 5) pay per view. While each of these has implications beyond

the sciences, it is in the sciences that they are today getting the most attention, so I will address each briefly.

### **7. Behavior modification**

Faculty councils and governance groups on numerous campuses have issued statements urging greater faculty awareness and making recommendations for appropriate and ethical behavior. These suggestions generally call for tenured faculty to forgo publishing in, reviewing for, or editing journals that do not adhere to certain standards (the standards being somewhat vague, but well understood). They recommend open access and professional society journals as preferred alternatives. Some faculty members are taking positive action, but the fact remains that many faculty members are relatively oblivious. For example, some 400 Triangle Area faculty members serve on editorial boards of Reed-Elsevier journals. In evaluating a journal, faculty generally do not consider whether it is a commercial journal, but rather submit to the highest quality journal they think might accept their work and edit for the most prestigious journals possible. Something stronger than moral guidance will be required to change this pattern, especially given that publication in commercial journals generally costs the author nothing, whereas open access journals and often those of professional societies charge to publish articles.

### **8. Open access**

Before the explosion of commercial journal outlets, scientists would usually pay page charges to professional society journals to publish their papers. Elimination of page charges and the transfer of costs to libraries was one of the mechanisms introduced early on by commercial publishers to lure manuscript submissions. The well-known consequence is that this limits the availability of our scholarly work to workers associated with rich libraries. A frequently discussed alternative is open access, which is essentially a return to a page-charge model but with the content distributed over the web, free of charge. With this model the author pays, rather than the libraries, making the information much more widely available. [PLoS](#), the most high profile experiment in open access, charges a flat fee of \$1500 per article and aspires to a profile equivalent to that of *Science* or *Nature*.

The economic consequences of general adoption open access have been widely discussed and widely distorted on both sides. For open access to work, resources for page charges need to be available to authors from research grants, new granting agency programs, publication funds of home institutions, or other sources. Authors have to be willing to spend the funds available to them instead of redirecting them toward their research programs. Publishers need to be able to price their open-access journals in such a way as to continue to make profits, either to support members services of professional organizations or reward share holders of commercial firms (albeit not necessarily at the 40% level Elsevier now enjoys). The commercial publishers have recognized the potential competition of open-access journals and in some cases now offer open-access publication in their otherwise expensive journals (e.g., Springer-Verlag\*\*). However, the economic incentives will favor authors submitting to low-cost publishers rather than those with high page charges, thereby putting competition in its proper place in the market and keeping library costs lower than at present. Academic institutions now face a complex situation rather like the classic Prisoner's Dilemma. If all institutions and faculty play along and move to open-access publication, they stand to save considerable money. However, individuals and institutions

can save even more by cheating and submitting to high-price commercial journals, and if too many cheat the system fails.

Must the open-access system hinge on altruism, or will there be incentives sufficient to lure authors to pay for open access for their articles? Last October Ted Freeman of Allen Press conducted an informal review of access rates as a function of age for articles published in the BioOne package of journals over the previous year and made a rather shocking discovery. For most traditional journals there is a pronounced peak in the first two months of availability, followed by a dramatic decline. For *Florida Entomologist*, a regional but long-established, open-access journal hit rates increased steadily over at least the first year following publication. By six months out, hit rates for articles in *Florida Entomologist* exceed those of the major international journal *BioScience*. We do not know how general or robust the pattern is, but it is suggestive of steady accretion of web linkages to open-access articles. Bergstrom & Bergstrom in a 2004 article in *Nature*\*\* report that, “Recent studies suggest that Open Access articles are cited much more often than similar articles without open access.” They go on to suggest that “Citations translate into both prestige and money; two recent econometric studies of economists’ salaries estimated that on average, controlling for age and number of articles published, doubling one’s number of citations increases one’s salary by 7-14%.” If open access is widely perceived to increase citation statistics, we may witness a natural transition. One model for easing this transition is that employed by the journal *Limnology and Oceanography*. *L&O* requires page charges for all articles, but for an additional fee (currently \$400) they will unlock an article and make it open access. These articles are more heavily used and cited. *L&O* reports that 199 of the top 200 articles downloaded in 2003 had been placed in free access\*\*. Thus, for the price of a set of old-fashioned reprints, an author can make an article open-access, be virtually assured of a larger readership, and have a high probability of increased citation levels.

### 9. Cutting out the middleman

A transition even more dramatic than the change to digital journal access has occurred in physics, mathematics, computer science and several related fields. Researchers now routinely post their manuscripts on preprint archives. As of January 2005, the granddaddy of digital archives, [arXiv](#), had received approximately 305,000 manuscripts during its 13-year history and was receiving new postings at a rate of approximately 3300 per month. Many alternative preprint archives have sprung up associated with specific disciplines or institutions. The advantages over traditional journals include immediate distribution and claim for priority, as well as dissemination not being constrained to rich libraries. The system is somewhat self-reviewing in that readers can generally post comments and authors can post updated versions, though the originals remain available more-or-less in perpetuity unless the author withdraws access to allow publication in a peer-reviewed journal. A parallel development in engineering and other technical fields is the rise of conference proceedings as the preferred initial form of communication, taking a role similar to that of the preprint servers, but often with rigorous quality review as a component of the selection process.

Most journals in fields where preprint servers or conference proceedings are used for initial dissemination and establishment of priority are pleased to consider these same contributions for subsequent publication, thereby providing a rigorous peer review mechanism that is decoupled from much of the priority and dissemination functionality. The long-term economic viability of

a publication system with decoupled peer review will hinge on the willingness of libraries to continue to support the somewhat redundant journals and the continued lack of other peer review options.

### **10. Ownership of copyright**

The University of North Carolina, like its peer institutions considers that its faculty own copyright to most of their creations. Thus, faculty are within their rights to give the copyright away, and commercial publishers generally insist on taking that copyright and letting the author or his institution retain few if any rights. However, universities would be within their rights to claim copyright to most of the work created by their faculty as works-for-hire, but one university doing so alone would place its faculty at an unfair competitive disadvantage. A university could strategically position itself for future action by asserting ownership and then readily giving it away at the request of faculty until such time as a large group of universities decided to change policy as a group.

Recently, government agencies have been examining the issue of ownership and distribution of articles derived from federally funded research. Proposals have been developed by Committees of the U.S. House of Representatives and the U.K. Parliament to encourage or require that articles derived from federally funded research grants be deposited in open archives after a short period of exclusive distribution by a scientific journal. In both the U.S. and the U.K. these proposals are controversial and far from settled. At this writing the U.K. proposal by the House of Commons Select Committee on Science and Technology has been blocked by a decision by the Department of Trade and Industry\*\*. The U.S. Congress approved on November 20, 2004 H.R.4818, the Consolidated Appropriations Act, wherein buried rather deeply was the observation that the period of request for comments had recently ended on an NIH proposal to request (not require) investigators to post on [PubMed Central](#) the final, peer-reviewed author's copies of manuscripts six months after the publication date, and requested that NIH give consideration to the comments and provide an estimate of implementation cost\*\*.

### **11. The long tail & pay-per-view**

A particularly unpopular but often recurring suggestion to ameliorate the library crisis is to shift from libraries providing universal access to at least a partial pay-per-view model. Despite the distaste this may conjure up among traditional library users, there are a number of compelling reasons to consider this as a solution. For example, consider that government boards that establish indirect cost rates are chronically resistant to inclusion of library expenses. If holders of federal grants were forced to include information access in their budgets, federal funding would become available through this alternative route and libraries might be able to cut back significantly on subscription expenses.

In a remarkably insightful article in *Wired Magazine*, Chris Anderson\*\* discussed the phenomenon of the long tail, or the large number of items accessed by only a few individuals, and how digital communication is a solution to the tyranny of space that precludes the tail in physical inventory. He observed that the average Barnes and Noble bookstore carries 130,000 titles, yet over half of Amazon.com's book sales come from outside the top 130,000 titles. While a typical music store will offer only a few tens of thousands of recordings, Rhapsody, a subscription-based streaming music service offers more than 735,000 tracks and almost all of

them are listened to by someone in a typical month. It streams more songs per month from outside its top 10,000 titles than within. The message here is that there is a demand for almost any content made available, and individual libraries can never hope to be sufficiently complete to come close to meeting that demand. Digital distribution provides everyone access to far more information than could possibly be provided in any reasonable physical space. We might extend this analysis to financial space. Libraries simply cannot afford to buy universal access to everything that anyone might want to use. It would likely be far more efficient to provide universal access to a core set of high-use resources, and then purchase on a pay-per-view basis access to additional content as needed. In this way the pool of information available to the scholar is vastly increased, and the library does not have to continually guess how to match today's collection development against tomorrow's needs.

## **12. Conclusions**

It is now seven years since Vince Resh wrote of peer-reviewed journal articles remaining the "coin of the realm" in the new era of digital communication, yet his statement remains accurate. Nonetheless, the landscape has changed. The transition to digital access and use is happening much more quickly than anticipated, and bringing with it new and richer opportunities for communication. In addition, digital access now assures that the individual scholar has access to much more information than ever imagined seven years ago, even if he or she has to pay for access. Dissemination and establishment of priority are to some extent becoming uncoupled from journal peer review with preprint servers and conference proceedings taking a progressively more central role. Digital access has also greatly increased the importance of citations because these links now provide a far more actively exploited mechanism for information discovery. Finally data registries, archives, and discovery tools are allowing heretofore impossible levels of information exchange, thereby enabling whole new areas of research and discovery.