Open Access: A New Paradigm for Knowledge Creation, Dissemination, and Access

Prepared by

William E. Moen
<William.Moen@unt.edu>
Director
Texas Center for Digital Knowledge
University of North Texas

Cathy Hartman
<Cathy.Hartman@unt.edu>
Assistant Dean of Libraries
University of North Texas

April 28, 2009
Open Access: An Unsettled Paradigm for Knowledge Creation, Dissemination, and Access

Introduction

The notion of open access to scholarly information is not new. In recent years, however, it has taken on prominence within the broader context of scholarly work, communication, and publishing. This brief paper intends to highlight and clarify key aspects of open access to assist UNT’s initial discussions of the utility of open access for UNT researchers and scholars.

The momentum for universities to consider open access policies has been growing and several major national research universities have approved policies related to open access. MIT is the most recent; its faculty voted in March 2009 to make their scholarly articles available to the public for free and to provide open access on the Web <http://web.mit.edu/newsoffice/2009/open-access-0320.html>. Stanford University’s School of Education adopted an open access policy in 2008 <http://ed.stanford.edu/suse/faculty/openaccess.html>. Three faculties at Harvard University (Faculty of Arts and Sciences, Law School, and Kennedy School of Government) voted in 2008 and 2009 to enact open access policies <http://osc.hul.harvard.edu/OpenAccess/policytexts.php>. In these cases, the universities’ faculty adopted and approved these policies.

The current interest in open access to scholarly information should be seen, however, in a longer time frame. Peter Suber created a timeline <http://www.earlham.edu/~peters/fos/timeline.htm> that captures what he considers landmark events in the open access movement, which he defines as:

the worldwide effort to provide free online access to scientific and scholarly research literature, especially peer-reviewed journal articles and their preprints.

His timeline is now being maintained at the Open Access Directory, which also provides links to a wide range of articles, papers, and other resources on the topic <http://oad.simmons.edu/oadwiki/Timeline> <http://oad.simmons.edu/oadwiki/Main_Page>.

The substantial increases in recent decades in the price of scholarly journals, sometimes called “the crisis in scholarly publishing,” have been a major motivation for making scholarly literature available through open access. UNT Libraries is spending $5.05 million this fiscal year on journals (both print versions and digital version via licensed databases). A discussion of that crisis and the broader issue of the economics of scholarly publishing, although an important issue, is outside the scope of this paper. Drott (2006) provides an introduction to some of the economic considerations of scholarly publishing and open access.

Scholarly Work, Communication, and Publishing

Any discussion of open access needs to be set in a broader context of scholarly work, scholarly communication, and scholarly publishing. Scholarly work comprises all the activities of researchers and scholars in pursuit of understanding, knowledge creation, and creative and artistic endeavors. Shaughnessy (1989) defines scholarly communication as “the social phenomenon whereby intellectual and creative activity is transmitted from one scholar to another”. He adds that scholarly communication is a “complicated process involving several important elements” as represented in Figure 1. Scholarly communication often includes the dissemination of the products of scholarly work in books, journals, and conferences.

Prior to the 17th Century, scholarly communication typically occurred through personal contact between scholars, either on an individual basis or at meetings of learned societies. Proceedings from these meetings were used to disseminate ideas, and the emergence of scholarly journals followed. *Philosophical Transactions of the Royal Society* and *Journal de Scavans* were the earliest peer-reviewed
journals, both first published in 1665. Peer-reviewed journals and conference proceedings continued to be primary dissemination and access mechanisms through the subsequent centuries until the Internet (Wells, 1999).

The opportunity for changing the dissemination and access mechanisms for scholarly information came with the advent of the Internet. While there were active efforts in using the Internet (e.g., through the use of listerv software and other tools) to produce online journals, it was the emergence of the Web (with the introduction of Mosaic, a graphical user interface browser in 1994), that presented even more threats to the traditional model of scholarly communication. The development and use of open source software in the form of digital repository software platforms (e.g., DSpace, Fedora Commons, and others) and open journal software sped the process of innovation and experimentation in providing access to scholarly information, in some cases bypassing traditional scholarly journal publishers.

Terminology

Before proceeding further, it may be helpful to provide definitions of key concepts related to open access.

Open access: Free availability via the public Internet to scholarly information, and permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the Internet itself. (paraphrased from the Budapest Open Access Initiative <http://www.soros.org/openaccess/read.shtml>).

Open access journals: Online journals (which may or may not be peer-reviewed) that provide articles in an open access manner. (The Directory of Open Access Journals <http://www.doaj.org/> lists over 4,000 such journals. See <http://ucblibraries.colorado.edu/dean/peer_reviewed.htm> for one list of peer-reviewed open access journals.)

Peer review: A specific form of quality review through which “a particular piece of research or scholarly communication is evaluated by the author’s peers [which might be the journal editor or external reviewers or both], who make a statement as to is value to researchers and the scholarly community. Peer review is designed to filter out the bad, the unimportant, or the uninteresting” (Weller, 2000, p. 1332).

Postprint: A manuscript that has been published in a peer-reviewed journal or conference proceeding.

Preprint: A version of a manuscript that has been peer-reviewed and accepted for publication.
Quality control: In the context of scholarly publishing, the processes and procedures whereby experts in a discipline review the contents of submitted manuscripts prior to publication to assure disciplinary standards are met for the quality of work represented in the manuscript.

Repository: A digital system that provides for the storage of digital information and access to that information (e.g., in the forms of preprints, postprints, technical reports, presentations, etc.).

Self-archiving: A process by which scholars and researchers deposit in an open access repository a copy (preprint or postprint) of a peer-reviewed journal article or conference paper.

Probably the most problematic term in the above list is “preprint”. Kling (2003, p. 598) states: “The preprint precedes a formally published version. An article should not be termed a preprint, however, until it has been accepted for publication in a specific venue.” He distinguishes a manuscript (or research manuscript) from a preprint, in that the former may be circulated by an author but has not yet been accepted for publication. Stevan Harnad and others focus on a definition of preprints that reflects a document that has been through a peer-review process AND accepted for publication. Traditional scholarly publishers may allow open access to articles in their journals – either preprints or postprints. With this in mind, we can offer definitions for two other terms that appear in the open access discussion:

Green open access: Refers to preprints and/or postprints (that have been peer-reviewed and accepted for publication or published) that authors self-archive for open access.

Gold open access: Refers primarily to open access journals but can include traditional journal publishers that provide open access to the published articles.

Open Access to Scholarly Publications

Stevan Harnad, one of the key figures in the open access movement, posted his “subversive proposal” in 1994 that proposed a new model for scholarly communication of “esoteric” (non-trade, no-market) scientific and scholarly publication (see Appendix A for a copy of the posting). In 1995, the Association for Research Libraries (ARL) produced Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing: An Internet Discussion about Scientific and Scholarly Journals and Their Future <http://www.arl.org/sc/subversive/>, a compendium of Internet discussions that Harnad’s posting stimulated. Ten years from his original post, Harnad posted The 1994 “Subversive Proposal for Electronic Publishing” at 10 in the American Scientist Open Access Forum <http://listserv.signumx.org/sc/wa.exe?A2=ind04&L=american-scientist-open-access-forum&D=1&O=D&F=I&S=&P=58204>. In this posting, Harnad clarified, refined, and revised some of the ideas in his original subversive proposal. His focus has been and continues to be open access to scholarly literature, specifically, peer-reviewed articles.

Over the years a number of events have moved the idea of open access forward. For example, there have been actions to mandate open access to scholarly literature. As one example, the National Institutes of Health (NIH) now mandate that publications that result from research funded by NIH be deposited in PubMed Central, an open access repository operated by the NIH <http://publicaccess.nih.gov/policy.htm>.

Legislation on open access has been debated in the U.S., specifically the Federal Research Public Access Act of 2006 introduced in 2006 by Senator John Cornyn (R-Texas) and co-sponsored by U.S. Sen. Joe Lieberman (D-Connecticut). The bill would require every federal agency with an annual research budget of more than $100 million to implement a public access policy, which would ensure that articles generated through research funded by that agency are made available online within six months of publication. One hundred and twenty-five provosts publicly supported this legislation <http://openaccess.eprints.org/index.php?/archives/135-125-_provosts_for,-10_against-frpaa-self-archiving-mandate.html>.
Open access mandates and initiatives are also occurring in other countries. In 2001, a meeting sponsored by the Open Society Institute resulted in the Budapest Open Access Initiative (<http://www.soros.org/openaccess/read.shtml>), which stated: “Removing access barriers to [open access] literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge”. The Registry of Open Access Repository Material Archiving Policies (ROARMAP) maintains a list of institutional and funding mandate policies for open access (<http://www.eprints.org/openaccess/policysignup/>).

Traditional scholarly journal publishers have also been addressing the open access concept by reconsidering their policies related to self-archiving. SHERPA (Securing a Hybrid Environment for Research Preservation and Access), funded by the U.K. Joint Information Systems Committee (JISC) and the Consortium of Research Libraries (CURL), has a project called RoMEO (<http://www.sherpa.ac.uk/romeo/>). This project compiles and maintains publisher's copyright and archiving policies (<http://romeo.eprints.org/publishers.html>). It categorizes publishers' policies as:

- **Green**: Publishers allow authors to archive preprint and postprint
- **Blue**: Publishers allow authors to archive postprint (i.e., final draft post-refereeing)
- **Yellow**: Publishers allow authors to archive manuscript prior to peer review and publication
- **White**: Archiving not formally supported by publisher.

Additional information about open access and scholarly publishing can be found from several sites, with the following being quite prominent:

- Steven Harnad’s Open Access Archivangelism blog: <http://openaccess.eprints.org/>
- Peter Suber’s The SPARC Open Access Newsletter <http://www.earlham.edu/~peters/fos/> and his Open Access blog <http://www.earlham.edu/~peters/fos/fosblog.html>

**Technology and Open Access**

Changing technology has provided opportunities for the open access movement. Most important have been the availability of digital repository platforms and the concept of “institutional repositories”. An institutional repository is a digital repository focused on collecting, disseminating, and preserving the intellectual output of an institution, typically an academic or research institution. According to statistics available from the Directory of Open Access Repositories (<http://www.opendoar.org>), there are nearly 1,200 institutional repositories operating worldwide. While a complete discussion about institutional and other digital repositories is beyond the scope of this paper, it is important to address these in the context of publishers’ open access policies.

Typically, when a publisher allows self-archiving, the article to be self-archived is required to be in an institutional repository. There are cases when a funding agency’s mandate prescribes a specific repository in which the publications are to be deposited (e.g., NIH requires the use of PubMed Central).

**Summary**

Although there are many aspects of open access to consider (e.g., impact of open access on citation rates), this paper has focused primarily on terminology and concept clarification. Its purpose is to establish an initial common ground for discussions at UNT on open access.
References


Appendix A: Stevan Harnad’s 1994 Subversive Proposal

We have heard many sanguine predictions about the demise of paper publishing, but life is short and the inevitable day still seems a long way off. This is a subversive proposal that could radically hasten that day. It is applicable only to ESOTERIC (non-trade, no-market) scientific and scholarly publication (but that is the lion’s share of the academic corpus anyway), namely, that body of work for which the author does not and never has expected to SELL the words. The scholarly author wants only to PUBLISH them, that is, to reach the eyes and minds of peers, fellow esoteric scientists and scholars the world over, so that they can build on one another’s contributions in that cumulative, collaborative enterprise called learned inquiry. For centuries, it was only out of reluctant necessity that authors of esoteric publications entered into the Faustian bargain of allowing a price-tag to be erected as a barrier between their work and its (tiny) intended readership, for that was the only way they could make their work public at all during the age when paper publication (and its substantial real expenses) was their only option.

But today there is another way, and that is PUBLIC FTP: If every esoteric author in the world this very day established a globally accessible local ftp archive for every piece of esoteric writing from this day forward, the long-heralded transition from paper Publication to purely electronic publication (of esoteric research) would follow suit almost immediately. This is already beginning to happen in the physics community, thanks to Paul Ginsparg’s HEP preprint network, with 20,000 users worldwide and 35,000 "hits" per day, and Paul Southworth’s CICnet is ready to help follow suit in other disciplines. The only two factors standing in the way of this outcome at this moment are (1) quality control (i.e., peer review and editing), which today happens to be implemented almost exclusively by paper publishers, and (2) the patina of paper publishing, which results from this monopoly on quality control. If all scholars’ preprints were universally available to all scholars by anonymous ftp (and gopher, and World-Wide Web, and the search/retrieval wonders of the future), NO scholar would ever consent to WITHDRAW any preprint of his from the public eye after the refereed version was accepted for paper "PUBLICation." Instead, everyone would, quite naturally, substitute the refereed, published reprint for the unrefereed preprint. Paper publishers will then either restructure themselves (with the cooperation of the scholarly community) so as to arrange for the much-reduced electronic-only page costs (which I estimate to be less than 25% of paper-page costs, contrary to the 75% figure that appears in most current publishers’ estimates) to be paid out of advance subsidies (from authors’ page charges, learned society dues, university publication budgets and/or governmental publication subsidies) or they will have to watch as the peer community spawns a brand new generation of electronic-only publishers who will.

The subversion will be complete, because the (esoteric -- no-market) peer-reviewed literature will have taken to the airwaves, where it always belonged, and those airwaves will be free (to the benefit of us all) because their true minimal expenses will be covered the optimal way for the unimpeded flow of esoteric knowledge to all: In advance.

Stevan Harnad
Cognitive Science Laboratory
Princeton University
Princeton NJ 08542
harnad@princeton.edu
June 27, 1994