

Open Access Self-Archiving of Refereed Research: A PostGutenberg Compromise

Stevan Harnad
Canada Research Chair in Cognitive Sciences
Université du Québec à Montréal
&
School of Electronics and Computer Science
University of Southampton
<http://openaccess.eprints.org>

If asked what would be ideal for their research access needs, most scientific and scholarly researchers would say that they would wish to have access to every piece of research relevant to their own work, rigorously peer-reviewed, conscientiously copyedited and elegantly formatted, online and on paper, as soon as it is ready for publication. (In some fields – e.g., high energy physics -- researchers also want access to research *before* it is peer-reviewed, but so far this is the exception rather than the rule.) Moreover, because in most fields the research users and the research authors are the same population, wearing different hats, what is ideal for the user is also ideal for the author: Researchers conduct and publish research so it can be accessed, used, applied and built upon by other researchers, in further ongoing research. The progress and funding of their scholarly work -- not to mention their careers and salaries -- depend on the uptake and impact of their research findings. Hence the broader and earlier the access to their findings, the better for authors (Gargouri et al 2010).

So much for ideals. Now, what is the reality? There are about 25,000 peer-reviewed scholarly and scientific journals, across all disciplines, nations and languages, publishing about 2.5 million articles per year. No university or research institution in the world can afford to subscribe to all, most or even many of those 25,000 journals; most can only afford to subscribe to a small fraction of them. That means that most researchers worldwide only have access to a small fraction of the research published annually; it also means that the authors of all those annual articles only have access to a fraction of their potential users worldwide. Access, usage, impact and research progress are being lost, annually, because access falls short of being universal.

A solution has existed ever since the onset of the PostGutenberg (online) era (Okerson & O'Donnell 1995). The solution is known, and it is (belatedly) beginning to be implemented: Authors can make their peer-reviewed research accessible free for all online by self-archiving their peer-reviewed final drafts in their institutional

repository immediately upon acceptance for publication, and their institutions and funders can mandate such self-archiving (Harnad et al 2003). The author's self-archived final draft is not the publisher's version of record – it is peer-reviewed but it is not copyedited nor in the publisher's final format. So the solution is a compromise; but it is a compromise that is incomparably better than the status quo. It means that refereed research findings are immediately available to all potential users, not just to the fraction that are at subscribing institutions. The published version's formatting is of no importance to the many would-be users who would otherwise have no access at all; and if the copyediting (which for most journals these days is exceedingly light¹) has corrected anything substantive, the author can update the final draft to incorporate that too.

Author self-archiving is called "Green Open Access" (Green OA). The majority of journals today (and almost all the top journals) have already given their official green light to immediate author self-archiving of their final drafts. For the minority of articles published in the journals that do not yet endorse Green OA, the final draft can and should be deposited in the author's institutional repository immediately upon acceptance for publication in any case. If the author wishes to observe a journal's embargo on OA, access to the deposit can be set as "Closed Access" rather than "Open Access" during the embargo. The bibliographic metadata (author, title, journal, abstract, etc.) of Closed Access deposits are immediately visible to all, worldwide, and the institutional repositories can implement an "eprint request" Button that allows would-be users to request and authors to provide a single copy for research purposes (Sale et al 2010). This too is a compromise: It is not OA; it is Almost-OA.

But universal Green OA self-archiving mandates, adopted by universities, research institutions and research funders worldwide will not only remedy the research access problem immediately, but it may eventually lead to an even better solution, and the natural one for the online era: Once the final drafts of all refereed research articles are being self-archived and hence freely accessible to all users online, institutions may well decide that they no longer need to subscribe to the journals in which they are published. Cancellation pressure will induce journals to cut costs by eliminating obsolete products and services, beginning with the print edition, and then the online edition. All production, access-provision and archiving will be offloaded onto the network of institutional repositories. The author's refereed, revised, accepted final draft, self-archived in his institutional repository, will become the version of record, and the only service still provided by the journal publisher will be peer review (and possibly some copyediting).

¹ Copyediting is the lightest in STM journals; it may still be somewhat more substantive in humanities and arts journals, as well as in books. This would need to be examined systematically but it seems almost certain that but the practise and the demand for copyediting is declining in the online era, and it may make more sense to offer it for a fee as ann optional extra service to authors and their institutions.

The true cost of peer review alone, per article, is only a fraction of what is being paid per article by institutional subscriptions today. Institutions will easily be able to cover the peer-review costs for their annual outgoing articles out of just a fraction of their annual windfall savings from the cancellation of their incoming journal subscriptions. That cost will be even lower if charged per individual round of refereeing as no-fault refereeing fees rather than as acceptance/publication fees (which require factoring in all the costs of the rejected articles into the fee for the accepted articles) (Harnad 2010).

Covering publication costs through per-article publication fees instead of through per-journal subscription fees is called "Gold OA publishing" (Harnad et al 2004). It is the natural, stable solution for refereed research publishing in the PostGutenberg era (Harnad 2009), but it is only possible if Green OA self-archiving is universally mandated first, so that (1) the access-provision and archiving costs can be offloaded onto institutional repositories, (2) the journals can downsize to peer-review service provision alone, and (3) institutional subscription cancellations can release the funds to pay for the peer review fees. Universal Green OA mandates followed by downsizing to Gold OA saves a good deal of money overall (Houghton et al 2009), whereas trying to do it the other way round costs more money and fails to generate universal OA (Harnad 2011).

Does this solution generalize to scholarly monographs? The economics of book publishing and journal publishing are not the same. Nor is it true of all authors of scholarly monographs, as it is true of all authors of peer-reviewed journal articles, that they right solely for uptake and impact, not for royalty revenue. But research is research, and book authors, too, benefit, both in their research and in their careers and funding, from the impact of their findings. So perhaps once a book citation index is created and shows the impact to be gained from making monographs OA, monographs too will take the Green and eventually the Gold road to OA (Harnad 2008).

Unlike with OA's primary target, journal articles, the deposit of the full-texts of books in [Open Access Repositories](#) cannot be mandated, only encouraged. However, the deposit of book metadata + plus + reference-lists can and should be mandated by universities and funders. That will create the metric that the book-based disciplines need most: a book citation index. Thompson-Reuters Web of Science only covers citations of books by (indexed) journal articles, but book-based disciplines' biggest need is *book-to-book citations*. [Citebase](#) could provide that, once the book reference metadata are being deposited in their authors' institutional repositories too, rather than just journal articles. (Google Books and Google Scholar are already providing a first approximation to a book citation count.) Analogues of "download" metrics for books are also potentially obtainable from book vendors, beginning with [Amazon Sales Rank](#). In the Humanities it also matters for credit and impact how much the non-academic (hence non-citing) public is reading their books ("Demotic Metrics"). Institutional repositories can not only (1) add book-metadata/reference deposit to their OA Deposit Mandates, but they can (2) harvest Amazon book-sales

metrics for their book metadata deposits, to add to their [IR stats](#). Repositories can also already harvest [Google Books](#) (and [Google Scholar](#)) book-citation counts today, as a first step toward constructing a distributed, universal OA book-citation index. The [Dublin humanities metrics conference](#) was also concerned about other kinds of online works, and how to measure and credit their impact: Metrics don't stop with citation counts and download counts. Among the many "Demotic metrics" that can also be counted are link-counts, tag-counts, blog-mentions, and web mentions. This applies to books/authors, as well as to data, to courseware and to other identifiable online resources. We should hasten the progress of book metrics, and that will in turn accelerate the growth in OA's primary target content: journal articles, as well as increasing support for institutional and funder OA Deposit Mandates.

REFERENCES

- Gargouri, Y., Hajjem, C., Lariviere, V., Gingras, Y., Brody, T., Carr, L. and Harnad, S. (2010) [Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research](#). *PLOS ONE* 5 (10). e13636
- Harnad, S. (2008) [Open Access Book-Impact and "Demotic" Metrics](#). *Open Access Archivangelism*. Friday, October 10, 2008.
- Harnad, S. (2009) [The PostGutenberg Open Access Journal](#). In: Cope, B. & Phillips, A (Eds.) *The Future of the Academic Journal*. Chandos.
- Harnad, S. (2010) [No-Fault Peer Review Charges: The Price of Selectivity Need Not Be Access Denied or Delayed](#). *D-Lib Magazine* 16 (7/8).
- Harnad, S. (2011) [Gold Open Access Publishing Must Not Be Allowed to Retard the Progress of Green Open Access Self-Archiving](#). *Logos* (in press)
- Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y., Oppenheim, C., Stamerjohanns, H., & Hilf, E. (2004) [The green and the gold roads to Open Access](#). *Nature Web Focus*.
- Harnad, S., Carr, L., Brody, T. & Oppenheim, C. (2003) [Mandated online RAE CVs Linked to University Eprint Archives](#). *Ariadne* 35.
- Houghton, J.W., Rasmussen, B., Sheehan, P.J., Oppenheim, C., Morris, A., Creaser, C., Greenwood, H., Summers, M. and Gourlay, A. (2009). [Economic Implications of Alternative Scholarly Publishing Models: Exploring the Costs and Benefits](#), London and Bristol: *The Joint Information Systems Committee* (JISC).
- Okerson, A. & O'Donnell, J. (Eds.) (1995) [Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing](#). Washington, DC., *Association of Research Libraries*, June 1995.

Sale, A., Couture, M., Rodrigues, E., Carr, L. and Harnad, S. (2010) [Open Access Mandates and the "Fair Dealing" Button](#). In: *Dynamic Fair Dealing: Creating Canadian Culture Online* (Rosemary J. Coombe & Darren Wershler, Eds.)