Open Access Self-Archiving of Refereed Research: A Post-Gutenberg Compromise

Stevan Harnad
University of Southampton, harnad@uqam.ca

Follow this and additional works at: http://docs.lib.purdue.edu/atg

Part of the Library and Information Science Commons

Recommended Citation
DOI: https://doi.org/10.7771/2380-176X.5786

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
practices to release material of the highest possible quality, published within known patterns and bearing standard mechanisms of cataloging and archiving, such as ISSSNs and DOIs. Journal article version metadata are now a facet of publishers’ responsibilities in disseminating scholarly material online. And, while SAGE cannot satisfy every researcher and every member of its community, SAGE is taking a significant step toward an industry-wide solution for standard versioning practices.

The next major hurdle in the evolution of journal article versioning is industry acceptance of post-publication corrections and enhanced versions of record. Many recommended standards, such as NISO’s JAV terms, incorporate support for any iterations following what was known in the print-only world as the “final” issue version or version of record. However, many publishers either do not make any changes to the version of record or display non-standard indicators when such changes occur. SAGE is prepared to contribute to shared efforts toward clear and acceptable practices for iterations beyond the version of record. SAGE is prepared to launch another wave of production and platform enhancements to our journals publishing program that allow clear indications of changes to an article’s version of record. We look forward to partnering with other members of the scholarly community to examine the conceptual and logistical implications of this change within to all aspects of our industry.

The 2010 NISO study supports this need for more discussion and awareness on these topics that will bring us closer to versioning standardization. Today, there exists a troubling disconnect between the needs of scholarly researchers and the obstacles encountered by those in publishing and dissemination roles. Further research of this sort is needed to expand our collective understanding of the type of demands from scholarly readers and practitioners for article versioning standards.

Open Access Self-Archiving of Refereed Research:
A Post-Gutenberg Compromise

by Stevan Harnad (Canada Research Chair in Cognitive Sciences, Université du Québec à Montréal & School of Electronics and Computer Science, University of Southampton) <harnad@ecs.soton.ac.uk> http://openaccess.eprints.org

I 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 copiededit, 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 Post-Gutenberg (online) era (Okersten and O’Donnell 1995). The solution is known, and it is (belatedly) beginning to be implemented: authors can make their peer-reviewed research accessible for all online by self-archiving their peer-reviewed final drafts in their institutional repositories 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 copyrighted 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 copyingediting (which for most journals these days is exceedingly light) 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, webwide, and the institutional repositories should 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
Unlike with OA’s primary target, journal articles, the deposit of the full-texts of books in Open Access Repositories (http://roar.eprints.org/) cannot be mandated (http://www.eprints.org/openaccess/policiesignup/), 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 (http://citebase.eprints.org/) 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 (http://www.rampant-books.com/mgt_amazon_sales_rank.htm). In the humanities it also matters for credit and impact how much the nonacademic (hence nonciting) public is reading their books (“Demitic Metrics”). Institutional repositories can not only (1) add book-meta-data/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 (http://rac.eprints.org/projects/irstats/). Repositories can also already harvest Google Books, http://books.google.com/books?q=%22decline+and+fall+of+the+roman+empire%22+gibbon&btnG=Search+Books, and Google Scholar, http://scholar.google.com/scholar?q=%22decline+and+fall+of+the+roman+empire%22+gibbon&sa=N&tab=ps) book-citation counts today, as a first step toward constructing a distributed, universal OA book-citation index. The Dublin Humanities Metrics Conference (http://www.coimbra-group.eu/DOUMENTS/coimbra-groups-seminars/metrics_workshop_programme2.pdf) 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 “Demitic 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


Endnotes

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 the practice and the demand for copyediting are declining in the online era, and it may make more sense to offer it for a fee as an optional extra service to authors and their institutions.