Title
Libraries Take on Policy: Support for Open Access and Open Data

Permalink
https://escholarship.org/uc/item/2hx264ms

Journal
Against the Grain, 26(2)

Author
Taylor, Anneliese S

Publication Date
2014-05-12

Anneliese Taylor

Libraries Take on Policy: Support for Open Access and Open Data

The growth of open access (OA) journal publishing has exploded in the last decade. The number of full, immediate OA articles went from 2% to 11% of all articles published between 2000 and 2011. When hybrid and embargoed open access articles are included in the count, the 2011 total jumps to 17% across all disciplines. And looking at biomedical journals specifically, 36% of articles published were OA[1].

Alongside this impressive growth in what's commonly referred to as gold open access publishing has been a multi-pronged effort to expand access to published articles through "green" open access. Green OA is the process of depositing a version of a published article in an open access repository, whether that be an institutional repository (IR) or a disciplinary repository, or even placing articles on an openly accessible web site. "Self-archiving" is frequently used interchangeably with green OA. It does not require authors to pay an article processing charge as many gold OA models do.

Many publishers have a history of allowing authors to self-archive a version of their article. The version is typically the accepted author's manuscript, incorporating changes from the peer-review process, but before the publisher has copyedited, formatted, and branded the manuscript for final publication. A very few publishers allow the final, published version to be uploaded via self-archiving. There may or may not be a delay period after publication before the manuscript can be made accessible.

Open access policies passed at the institutional level or by research funders are an attempt to broaden public access on a larger scale. The potential to open up access via a formalized policy is significant, but not without some effort. The next sections will highlight health sciences libraries and their roles with green OA policies in the United States.

NIH Public Access Policy

The single largest influencer on the growth of green OA articles in the health sciences to date is the National Institutes of Health (NIH) Public Access Policy, passed in 2008. This policy requires peer-reviewed scholarly articles published by NIH-funded researchers to be submitted to PubMed Central (PMC). The articles are made publicly available no later than twelve months after the official publication date.[2] PMC is NIH's freely accessible, full text article repository, with close to 3 million articles currently deposited. NIH Policy manuscripts are about 10% of this amount. The rest of the content comes from publishers that voluntarily deposit their articles, usually after a publication delay.

Many health sciences (HS) libraries got involved in 2008 or earlier with support for researchers at their academy needing to comply with the policy. Efforts range from offering workshops and individual consultations to notifying authors which of their articles are non-compliant and what steps to take to make them compliant. There are a number of excellent NIH Policy LibGuides (e.g. Duke University and University of Washington) and video tutorials (e.g. New York University and Harvard University) created by librarians to assist researchers.

An informal survey in 2013 of Association of Academic Health Sciences Libraries (AAHSL) member involvement in supporting the Policy indicated a high level of activity amongst health sciences libraries. Of the 25 responses, all but four were actively involved with a support role. It's remarkable that HS

libraries are taking on this new responsibility at a time with shrinking staff and budgets. The work can be very involved and time-consuming, so it is no small decision to take it on.

Here are a few notable initiatives:

- University of Arkansas for Medical Sciences Library runs reports on the NIH Public Access Compliance Monitor (PACM) of UAMS authors' adherence to the Policy on a departmental, institutional, and individual PI level. The library shares the documents it uses to notify authors for other institutions' benefit[3].
- The Countway Medical Library at Harvard University built an online submission system where authors can deposit their manuscripts. A librarian logs in to the NIH Manuscript Submission System (NIHMS) as a publisher, enabling deposit of multiple papers on behalf of authors[4]. After a coordinated outreach effort, Harvard's compliance rate jumped to the 90% range.
- With help from a grant, Health Sciences Libraries staff at New York University programmed an automated system to notify School of Medicine authors who have published articles that are not compliant with the policy. The program matches PACM data with an internal Sponsored Programs Administration database to identify active grants and contact emails. Seven months after the library starting sending monthly email notices in June 2013, the SOM's compliance rate rose from 79% to 87%.[5]
- On behalf of authors with non-compliant articles published in "Method D" journals (where publisher makes initial deposit), the University of California, San Francisco Library sent lists to several publishers with a request to deposit the manuscripts into NIHMS. Results were mixed but fortunately the publisher with the most non-deposited articles agreed to deposit all manuscripts.

The advantages to the library are the opportunity to collaborate with different groups on campus, and filling a needed role. Offices of sponsored research have welcomed libraries' help in supporting researchers trying to work through the complications of getting the PMCID, which indicates an article is compliant. Efforts where research offices and the library work as a team are particularly effective, as each group has its own strength. Librarians have established relationships with publishers and are accustomed to reading contract agreements and to finding the needle in the haystack. There are several steps to compliance with the NIH Policy, leaving plenty of room for error, so librarians' perseverance with detail comes in handy.

Librarians who have discussed this topic informally among themselves often indicate that benefits include the opportunity to do outreach in new ways and with new constituents at their organization as benefits. Putting librarians in a public service role related to the policy is good exposure for the library, so long as the library comes off as a helpful resource and not simply as enforcement. The level of support is determined by the library's priorities and available resources, by institutional culture, and by acceptance from campus groups of the library taking the lead in research policy compliance support.

You might be wondering, is it worth it for the library (or anyone else for that matter) to go to all of this effort? Consider the fact that PMC gets over 700,000 unique visitors daily, from around the world (a fact that doesn't escape journal publishers[6]). Studies conducted in 2010 and 2011 have found that health personnel read primary literature found in PubMed and that access contributes to health care[7, 8].

Expanding Public Access to Federally-Funded Research

The need for library support for the NIH Public Access Policy is likely to lessen over the next few years as authors get continually comfortable with the routine. Meanwhile, a similar policy will be expanding to other U.S. federal agencies. Under President Obama, the White House Office of Science & Technology Policy (OSTP) issued a directive in February 2013 to all federal agencies with more than $100M in research & development expenditures to develop a public access policy. It calls for making the direct results of federally funded research, both peer-reviewed articles and digital data, publicly available and useful.[9]

The directive extends a public access policy to around 20 agencies, including the National Science Foundation and the Department of Health and Human Services. Another interesting part of this order is the inclusion of data and metadata and the focus on the usability and preservation of research outputs. Under the NIH Public Access Policy, data are not included, and access is the only thing specified. The OSTP highlights the importance of being able to search, retrieve, and analyze data in digital formats to enable scientific breakthroughs and stimulate innovations.

On the data front, researchers will be required to develop data management plans and will be expected to deposit data in publicly accessible databases "where appropriate and available".[9] Since some publishers (PLOS, Nature) and some funders such as the National Science Foundation and the NIH already require data management plans, several academic libraries have already established programs around data management and data curation. These programs are multi-disciplinary and therefore not necessarily based in the health sciences library, however HS subject liaisons are increasingly involved in helping researchers make plans to properly store, preserve, and share their data.

Data is the new currency for research[10], and libraries and their institutions are increasingly partnering to develop data curation infrastructure and services. The DataConservancy project is the outcome of a $20M NSF grant awarded to the Johns Hopkins University Sheridan Libraries. The project provides tools to preserve, share, and discover data. A popular, free tool for generating data management plans is the DMPTool, run by the California Digital Library (CDL).

The datasets themselves may be deposited in a growing selection of open data repositories, which are managed by government agencies, non-profit and commercial organizations, and academic institutions. The UCSF Library partnered with the UCSF Clinical & Translational Science Institute and the CDL to develop an open data repository for UCSF scientists called DataShare, released in 2013. Data is preserved in CDL's Merritt repository and each dataset is assigned a unique EZID identifier for tracking and citation. Content is gradually being added, as researchers overcome the hurdle of discomfort with releasing datasets too soon.

It remains to be seen what kind of an impact the OSTP directive will have on libraries, as the office has yet to release plans for carrying out the policy. The OSTP specifically stated its preference for agencies to work together in developing their plans, and to leverage existing archives. It also encourages public-private partnerships where appropriate.

If the chosen model is along the lines of what the publisher-backed CHORUS has proposed, the content would reside on publishers' servers and would require less involvement by authors and librarians who support them (though it's hard to imagine this system being fool-proof; then there's the question of

compliance verification). If the SHARE model is adopted, which relies on university-based digital repositories, then library involvement will be significant, as the stewards for institutional repositories. Of course, the government might opt for more than one method, depending on the agency. Then we'll certainly have our hands full!

Organizational Open Access Policies

The U.S. federal research funders have taken the green OA path toward providing access to publicly-funded research. This differs from the Research Councils U.K.'s (RCUK), which sets aside block grants to cover article processing charges for paid, gold OA, but allows embargoed, green OA when gold OA is not an option.

Yet another approach to broadening access to scholarly research is the institutional open access policy. There are currently 250 institutional or sub-institutional OA mandates in place around the world, and the list grows monthly. The policies are passed by academic faculty or researchers, and libraries take on the implementation. Librarians in all disciplines have become deeply engaged in the intricacies of these policies and in supporting authors depositing their final manuscripts in the institutional repository.

Institutional OA policies have a great potential to expand (true) open access (not just public access) to a vast quantity of peer-reviewed scholarly articles. In practice, however, the compliance rate is very low as it relies on authors voluntarily depositing their articles in an open repository – without the threat of losing funding if they don't take the steps. Always the information organizers, several libraries have invested in database solutions to manage author publication tracking and deposit workflows. The libraries at Duke University, MIT, and the University of California have all initiated implementation of such a tool (managed by but not necessarily funded by the library). The effectiveness of these projects will be evaluated after a period of implementation and will inform the library community at large about what kind of results can be expected as a return on the library's commitment to supporting OA policies.

Looking Ahead

There is no doubt that the library's role within the academy is going through significant transition, especially when it comes to health sciences fields. As the nature of clinical practice, research, and scholarly communication changes with technological advances, libraries are finding their niche and trying on new roles that build on librarian strengths.

References

1. Mikael Laakso and Bo-Christer Björk, "Anatomy of open access publishing: a study of longitudinal development and internal structure," *BMC Medicine* v.10#124 (2012), doi:10.1186/1741-7015-10-124.
2. NIH Public Access Policy http://publicaccess.nih.gov/index.htm
3. University of Arkansas Medical Sciences Library. "NIH PAP". Accessed January 13, 2014, https://docs.google.com/folderview?id=0B6-FIjB8QQKHUV9SOWpWT0h4Q2c

4. P. Scott Lapinski, David Osterbur, Joshua Parker, and Alexa T. McCray, "Supporting Public Access to Research Results," *College & Research Libraries* v.75 (January 2014):20-33.

5. Stephen Maher, telephone conversation with author, January 9, 2014.

6. Philip M. Davis, "Public accessibility of biomedical articles from PubMed Central reduces journal readership--retrospective cohort analysis," *FASEB Journal* v.27#7:2536-41 (July 2013), doi:10.1096/fj.13-229922.
7. Jamie O'Keeffe, John Willinsky, Lauren Maggio, "Public access and use of health research: an exploratory study of the National Institutes of Health (NIH) Public Access Policy using interviews and surveys of health personnel," *Journal of Medical Internet Research* 13#4: e97 (2011), doi: 10.2196/jmir.1827.

8. Lauren A. Maggio, Ryan M. Steinberg, Laura Moorhead, Bridget O'Brien, John Willinsky, "Access of primary and secondary literature by health personnel in an academic health center: implications for open access," *Journal of the Medical Library Association* 101#3:205-12 (July 2013), doi: 10.3163/1536-5050.101.3.010.
9. White House Office of Science and Technology Policy, "Expanding Public Access to the Results of Federally Funded Research." Accessed January 13, 2014, http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research.
10. John Markoff, "How to Share Scientific Data," *New York Times*, August 12, 2013, http://nyti.ms/1aPQSTx.