

2014

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

Anneliese Taylor

*Scholarly Communications and Collections*, [anneliese.taylor@ucsf.edu](mailto:anneliese.taylor@ucsf.edu)

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



Part of the [Library and Information Science Commons](#)

---

### Recommended Citation

Taylor, Anneliese (2017) "Libraries Take on Policy: Support for Open Access and Open Data," *Against the Grain*: Vol. 26: Iss. 2, Article 11.

DOI: <https://doi.org/10.7771/2380-176X.6696>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

**Redundancy** — Busy researchers in academic health centers — many of whom are required to submit manuscripts to **PubMed Central** — often perceive self-archiving in institutional repositories to be a redundant activity. Some can be convinced by arguments that the library is building a comprehensive collection of university scholarship, including faculty papers that cannot be made available through **PubMed Central**, and that it is important not to rely on external sources to make these publications accessible. This is an ongoing issue for the administrators of medical institutional repositories to address, and in fact, some have decided not to routinely collect papers from **PubMed Central**. Implementing automated harvesting and direct deposit solutions that reduce researcher effort are critical. A federated network of repositories as envisioned by SHARE may lessen or even eliminate redundancy in the future.

**Staffing, Sustainability, and Scalability** — Like other academic libraries, health sciences libraries have employed various approaches for staffing and allocating resources for their institutional repositories. Opportunities for medical libraries tend to be more limited, since in general they have smaller staffs and may not have access to undergraduate student labor to perform repository tasks. The staffing dilemma may lead medical libraries to rely on the institutional repositories managed by the main campus library, or make them hesitant to become actively involved.

Some medical libraries have moved forward by repurposing or reprioritizing existing librarians and paraprofessionals from areas

where services have declined, such as cataloging or circulation. Cross-departmental teams are also an option. Permanent staffing and strong leadership clearly help with repository promotion, content recruitment, and building trust and credibility. Departmental administrative staff can be enlisted for assisting with deposits for their departments. Technology and automated solutions should be explored.

Staffing for repository work is certainly a challenge, but investing in this work allows the library to provide a valuable service to the community and builds relationships among library staff and researchers. Support of library administration for institutional repositories is a key factor for successful and creative staffing solutions.

### Looking Ahead

Academic libraries are confronting rapid changes in higher education and scholarly communication. As open access and research data sharing gain momentum, institutional repositories have taken root at many academic libraries, including health sciences libraries, and are becoming a critical component of the services that libraries provide to their researchers, faculty, staff, and students. The care, encouragement, and patience of repository administrators are paying off, and the utilization of institutional repositories is growing to include publishing, grant support, and the measurement of research impact. The promise and potential of a federated network of repositories are compelling. Medical libraries should continue to cultivate their institutional repositories, which in turn allow them to cultivate and disseminate scholarship produced at their institutions. By playing a critical leadership role in this area, medical libraries can gain visibility and credibility across the institution, expand the skills

and expertise of library staff, and build new partnerships and collaborations.

### Acknowledgment

This paper is based in part on ideas presented in a 2011 Webinar, “Challenges and Opportunities for Medical Institutional Repositories” with **Dan Kipnis** and **Ann Koopman**, and a presentation at the **Special Libraries Association Annual Conference**, Chicago, Illinois, July 2012. 🌿

### Endnotes

1. **Clifford A. Lynch**, “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age.” *ARL* no. 226 (February 2003): 1–7, accessed December 21, 2013, <http://www.arl.org/storage/documents/publications/arl-br-226.pdf>.
2. **University of Nottingham**, “OpenDOAR, the Directory of Open Access Repositories,” last modified December 20, 2013, accessed December 21, 2013, <http://www.openoar.org/index.html>.
3. **Association of Academic Health Sciences Libraries**, “Annual Statistics of Medical School Libraries in the United States and Canada, Services 2010,” accessed December 30, 2013, <http://www.aahsl.org/annual-statistics>.
4. **Bepress**, “Medical School Institutional Repositories,” accessed December 21, 2013, <http://digitalcommons.bepress.com/institutional-repository-medical/>.
5. **Confederation of Open Access Repositories**, “Incentives, Integration, and Mediation: Sustainable Practices for Populating Repositories,” June 2013, accessed December 21, 2013, <https://www.coar-repositories.org/activities/repository-content/sustainable-practices-for-populating-repositories-report/>.

---

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

by **Anneliese Taylor** (Assistant Director for Scholarly Communications & Collections) <[anneliese.taylor@ucsf.edu](mailto:anneliese.taylor@ucsf.edu)>

---

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.<sup>1</sup>

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 Website. “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

*continued on page 30*

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.<sup>2</sup> **PMC** is **NIH's** freely accessible, full text article repository, with close to three 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<sup>3</sup>.
- 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.<sup>4</sup> 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

## against the grain people profile

Assistant Director for Scholarly Communications & Collections  
University of California, San Francisco Library  
530 Parnassus Avenue, San Francisco, CA 94143-0840  
Phone: (415) 476-8415 • <anneliese.taylor@ucsf.edu>

### Anneliese Taylor

**BORN AND LIVED:** I grew up in Houston, TX and attended **Sarah Lawrence College** in Bronxville, NY. During college I spent a summer studying French at **Middlebury College** in beautiful Vermont, and a year abroad in Paris and Caen, France. After college I moved to Austin, TX where I got my MLIS, and since then I have lived in Washington, DC, Philadelphia, and now San Francisco.

**PROFESSIONAL CAREER AND ACTIVITIES:** Though I'm naturally attracted to the social sciences, I've found my home working in STEM libraries. My first librarian position was as Engineering Liaison Librarian at **George Mason University** in Fairfax, VA. I then worked as a Science Librarian at **Bryn Mawr College** in Pennsylvania. In 2003 I took the plunge and moved across the country for a job as Collection Development Manager at the **University of California, San Francisco**, a health sciences university. I've been incredibly impressed by the brilliant, accomplished, yet down-to-earth students and faculty I've worked with. I'm also psyched to be working on scholarly publishing and communications initiatives.

**IN MY SPARE TIME:** These days I enjoy anything that gets me away from a computing device.

**FAVORITE BOOKS:** Recently — the *Millennium Trilogy* by **Stieg Larsson**.

**PET PEEVES:** Wasting time.

**MOST MEMORABLE CAREER ACHIEVEMENT:** Has been the opportunity to transition to a role as supervisor and manager in my job at **UCSF**. It's a honor to be given the responsibility, and has helped me grow both personally and professionally.

**HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS:** I see libraries collaborating more and more with groups outside the library on our campuses and within our institutions, to build partnerships and expand our portfolios in new ways. I think this is essential to demonstrating the need for librarians. I also see libraries having a better understanding of publishing and taking on a publishing role itself, as well as influencing tools used for scholarly communication. 🌱



the library started sending monthly email notices in June 2013, the **SOM's** compliance rate rose from 79% to 87%.<sup>5</sup>

- 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 es-

continued on page 32

cape journal publishers<sup>6</sup>). Studies conducted in 2010 and 2011 have found that health personnel read primary literature found in PubMed and that access contributes to health care.<sup>7,8</sup>

### 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.

On January 17, **President Obama** signed the 2014 omnibus appropriations legislation, thereby codifying a portion of the **OSTP** directive.<sup>9</sup> The new law calls for the manuscripts of articles funded through awards from the Departments of Education, Health & Human Services, and Labor to be made publicly accessible no later than 12 months after publication. Other parts of the directive are not addressed in the legislation and information has yet to come on how these aspects will be addressed.

The directive extends a public access policy to around 20 agencies, including the **National Science Foundation** and the **Department of Energy**. 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.” 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,<sup>10</sup> 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 any system not requiring some level of intervention and 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

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. 🌱

### Endnotes

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://public-access.nih.gov/index.htm>.
3. **University of Arkansas Medical Sciences Library**. “NIH PAP.” Accessed January 13, 2014, <https://docs.google.com/folderview?id=0B6-FIjB8QQKHUV9SOW-pWT0h4Q2c>.
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. **Ian Chant**, “Federal Spending Bill Expands Research Funding with Open Access Mandate, Restores IMLS Funding,” *LibraryJournal*, January 23, 2014, <http://bit.ly/1nUNH6C>.
10. **John Markoff**, “How to Share Scientific Data,” *New York Times*, August 12, 2013, <http://nyti.ms/1aPQSTx>.