



The Political Economy of Federally Sponsored Data

Bart Ragon

University of Virginia, Charlottesville, VA, USA

Abstract

Librarian involvement in the Open Access (OA) movement has traditionally focused on access to scholarly publications. Recent actions by the White House have focused attention on access on the data produced from federally sponsored research. Questions have emerged concerning access to the output of federally sponsored research and whether it is a public or private good. Understanding the political battle over access to federally funded research is closely tied to the ownership of the peer review process in higher education and associated revenue

streams, and as a result, interest groups seeking to influence government regulation have politicized the issues. As a major funder of research in higher education, policies from the federal government are likely to drive change in research practices at higher-education institutions and impact library services. The political economy of federally sponsored research data will shape research enterprises in higher education and inspire a number of new services distributed throughout the research life cycle.

Introduction

The nature of science is changing due to the evolution of networked technologies. Collaborative networked science is dramatically impacting scientific research in higher education and shifting norms of access, inspiring several movements which seek to “open” knowledge and resources using network technologies. Hey, Tansley, and Tolle refer to these recent evolutions in science as the Fourth Paradigm, “science based on data-intensive computing” (Hey, Tansley, and Tolle 2009). Librarians have supported this paradigm shift by becoming involved in eScience -- focusing on data, data management, and research support services along the research life cycle. Independently, librarian involvement in the Open Access (OA) movement has focused on access to scholarly lit-

erature – the synthesized output of the research process. Much of this interest can be attributed to declining library budgets and increased journal subscription costs; however, for many, OA is an ideological belief that access to knowledge is the cornerstone of our educational enterprises. This ideology is transferring to the evolving discussion on the political stage regarding access to the research data as it becomes increasingly important to the advancement and reproducibility of science.

Conceptually, OA includes ancillary concepts like Open Science and Open Data and library involvement, which can be seen in both of these movements. The focus of the debate is whether access to the output of federally funded research is a public or private good, and as such the issues have be-

Correspondence to Bart Ragon: bart@virginia.edu

Keywords: data services, libraries

come politicized. Public goods can be thought of as products collectively available to society where “one person’s use doesn’t diminish the supply for everyone else” (Stone 2012). Universities are often associated as a public good and have been called the producers of universal knowledge and information (*Universities and the Public Sphere: Knowledge Creation and State Building in the Era of Globalization* / 2012). These concepts collide over access to publications when produced from federally funded research but published in for-profit journals. Peter Suber, director of the Harvard Open Access Project, asserts that:

“There is a very important role for government in assuring open access. One reason there should be government involvement is that most scientific research, not most humanities research, but most scientific research is funded by taxpayers. It’s publically funded. There are several ways in which we’ve already paid for this research and if we’ve already paid for it, then we deserve access to it. It’s partly a matter of simple fairness, but is also a matter of getting the best bang for our buck, or getting the greatest return on taxpayer investment in this research” (Hagenmaier, Rolando, and Rascoe 2013).

The relevancy of OA in the realm of data is evident in a slight shift in language for National Science Foundation (NSF) biosketch requirements. In January 2013, language in NSF Biosketch requirements was changed from “Publications” to “Products” including products other than peer-reviewed papers, such as data sets, software, patents, and copyrights, to be reported on the grant application (“NSF Broadens Biosketch Publications Category to Include Products; Society Recommends Similar Change to NIH” 2013). How will the practice of eScience librarianship be impacted as federal requirements adjust to include access to data? More importantly, interest groups on either side of the issue seek to influence policy or legislation from governing bodies regarding OA, yet their concerns over access to all products of federally funded research vary. Regardless, these interest groups are impacting the con-

versation on the national stage and their efforts and perspectives should be monitored closely by librarians for impact on their institutions. Library literature on OA often focuses on the academic perspective, providing insight to the issues from within the context of higher-education institutions. However, this does not account for political perspectives, policies, and legislation, which potentially can have more influence on the practice of grant-funded research within higher-education institutions. Through policy changes and legislation, the government is acting on what it perceives as the best interest of the polis. In this capacity, the government acts as an agent of the State in its stewardship of tax payer money the federal government has the authority to establish, and enforce policy and legislation around products to federally funded projects. The perspective of the federal government and interest groups who seek to influence policy must be considered to understand the scope and direction of this complex and evolving process and how it will impact future access to data.

Interest Groups

Interest-group theory provides a useful lens in understanding the past and current political contest over the OA movement. In order to understand interest group theory it is first important to differentiate an organized interest group from groups of interested individuals. Educational literature lacks a unified definition, potentially making the distinction ambiguous (Bruce Cooper, James Cibulka 2008). Individuals on either side of the OA debate represent the interests of authors, libraries, publishers, taxpayers, and government agents. Interest-group formation is a dynamic process, where “leaders define an interest by portraying an issue, showing how it affects people, and persuading them that the portrait is accurate; and leaders speak for people in the sense of standing for them and articulating their wishes in policy debates” (Stone 2012). OA is part of broader ‘open’ social movements, enabled by networked-based technologies and a belief that

shared information will lead to shared knowledge. While previous focus has been on publications, advancements in the same network technologies are generating similar questions about access to data as a knowledge source. Elements of the OA social movement have been consolidated into the formation of organized interest groups. Lee (1969) alludes to this transformation as the, "crystallization of much of the motivation of dissatisfaction, hope, and desire awakened by the general social movement and focusing of this motivation on some specific objective." Crystallization of the interest groups around OA has been driven by questions pertaining to intellectual property, copyright, and revenue streams since the late 1990s, when the National Institute of Health (NIH) sought to establish a repository for biomedical literature from federally funded research. In the battle over federally funded research products, it is important to understand that there are groups of individuals who share common interests, but policies and mandates are far more likely to affect practice. For example, the 2005 NIH public access policy requesting grantees to submit final manuscripts to PubMed Central yielded low compliance rates; however, compliance increased to around 75% when the policy became mandatory in 2008 (Matthews 2013). This impact can be seen in the historical context of governmental regulation around NIH public access policies. More comprehensive documentation can be found on the *Timeline of the Open Access Movement* which provides detailed information from the 1960s – 2009 about the Open Access movement and the complexity of action-reactions created interest groups ("Peter Suber, Open-Access Timeline (formerly: FOS Timeline)").

Questions concerning the intellectual property for the products of federally funded research mobilized political interest groups in the OA Movement. Interest groups in support of OA to federally funded research include the Scholarly Publishing & Academic Resources Coalition (SPARC) and Alliance

for Taxpayer Access (ATA). These groups represent professional library organizations, universities, academic and research libraries, patient groups, physicians, researchers, publishers, and health promotion organizations. SPARC argues that, "U.S. taxpayers underwrite this research; they have a right to expect that its dissemination and use will be maximized, and that they themselves will have access to it. If this information is shared with all potential users, it will advance science and improve the lives and welfare of people of the United States and the world" ("Federal Research Public Access Act (SPARC2)" 2013). In opposition to additional regulation, the Association of American Publishers (AAP) represents the interests of commercial, educational, and professional companies, non-profits, university presses, and scholarly societies. Publishers report investing \$1,500 to \$4,000 for each research article published and maintain that they should be able to expect a reasonable return on their investment ("Overview of STM Publishing Value to Research" 2008). In many cases, publishers make content freely available after an embargo period that allows them to generate revenue in the market. However, the issues around OA are not so black and white. As Bosch et al. (2011) state, "Libraries do not have the resources to continue to exist in a world of ever-increasing prices, nor can publishers exist in world of no revenue increases."

SPARC and the AAP each produced position statements on OA. SPARC asserts that OA benefits researchers, educational institutions, businesses, the public, and funders. According to SPARC, "Funders invest in research in order to accelerate the pace of scientific discovery, encourage innovation, enrich education, and stimulate the economy – to improve the public good" ("Why Open Access?" 2013). They go on to claim that OA carries many benefits including increased relevancy, discoverability, democratization of information, increased return on investment, and enabling interdisciplinary research. SPARC believes transparency of information

will broaden the community of potential users and that if research is funded by taxpayers, it should be available to the public. Conversely, the position of the AAP is that publishers provide the maximum benefit to the public by enforcing the peer-review process in scholarly literature. The AAP states, "The public benefits from the scientific publishing industry producing organized, high quality, peer-reviewed journals" ("The Association of American Publishers - Open Access" 2013). They argue that if researchers are forced to publish in OA journals it limits choice in disseminating findings. They also point out that OA journals have a different business model and that publishing in an OA journal could shift the burden of cost to the researcher or their institution.

Battles between interest groups over OA have continued through a sequence of legislative and policy efforts since 2005. For example, the Federal Research Public Access Act (FRPAA) introduced in 2006, 2009, and 2012 (though never passed) provides the context to observe actions and reactions from interest groups from both sides of the issue. FRPAA sought to expand the NIH policy to 11 other major federal funding agencies including the National Science Foundation, Department of Energy, and Department of Defense. SPARC provides advocacy information specifically targeted for higher-education campuses, which often serve as incubators for activism in social movements, including information about FRPAA and links to supporting information from the Alliance for Taxpayer Access as well as means for individual and organizational advocacy. SPARC's approach was to increase mobilization by providing a call to action where individuals and leaders in higher education can engage as supporters of the bill. The AAP strategy is different than the approach by SPARC. AAP hired Eric Dezenhall, a "pit-bull of PR," who bypassed the consumer audience entirely (Dyer 2007). Dezenhall recognized that when given the choice, arguments that speak to free access are likely to resonate with taxpayers. The

AAP worked directly with decision makers arguing that, "public access equals government censorship" and "scientific journals preserve the quality/pedigree of science" (Dyer, 2007). Their campaign included writing letters to congress, NIH, and the White House Office of Science Technology Policy (OSTP). They also filed lawsuits and developed position statements representing the interest of the publishers they represent. The AAP portrayed themselves as a special interest group working towards the public good, achieving some success with the introduction of the Fair Copyright in Research Works Act of 2009 by John Conyers (D-MI) and Research Works Act of 2011 by Darrell Issa (R-CA). Both bills sought to abolish federal policies requiring mandatory deposit for content produced in the private sector, though neither passed.

Enter Access to Data

In February 2013, two key issues emerged reigniting the political contest over OA. First, the White House OSTP released a new policy requiring agencies to ensure access to federally funded publications and digital data are available and discoverable to the public (Stebbins, n.d.). Second, the Fair Access to Science and Technology Research Act (FASTR) was introduced in both the House and Senate and would require agencies to ensure public access to publications from federally funded projects. While the OSTP policy and FASTR are similar in their intent, there are some key distinctions. Notably, the OSTP policy specifically addresses access to digital data while FASTR focuses only on publications. Librarians should view each of these efforts by the government not only with the differences in their wording, but for their strengths and weaknesses as political devices. The OSTP policy is far more flexible, leaving means for adjustments by individual agencies. This opens the possibility of influence from interest groups. Additionally, FASTR would require that authors deposit in freely available public repositories, as opposed to publishers providing free ac-

cess after an embargo period. SPARC supports both the OSTP policy and FASTR legislation as efforts to increase OA in the public sector. AAP supports the OSTP policy and maintains that requiring federal agencies to duplicate already established resources and systems wastes taxpayer money (Sporkin 2013). AAP claims that FASTR undermines current efforts of publishers working with partners in higher education to provide OA to publications. AAP asserts that it is a partner with federal funding agencies, universities, and investigators in providing public access to published findings as part of the peer-review process.

The difference between a policy and legislation is also important when assessing the potential impact on access to data. As a White House directive, the OSTP policy can have a more immediate impact, but could be overturned by subsequent administrations. Unlike the OSTP policy, if passed, FASTR would not be as susceptible to changes in OA ideology by future administrations. In September 2013 the Public Access to Public Science Act (PAPS) was introduced in an attempt to codify language from the OSTP policy into legislation, however this legislation was viewed as weaker by SPARC (“New Public Access Legislation Introduced in U.S. House of Representatives | SPARC”).

Responses from Interest Groups

Within six months of the OSTP memorandum, February 22, 2013, federal agencies with more than \$100 million supporting research and development were directed to develop a draft plan to support increased access to the results of federally funded research (Holdren 2013). Since the memorandum, interest groups have realigned in an effort to influence final policy decisions. The AAP sponsored Clearinghouse for the Open Research of the United States (CHORUS) is a proposal from a coalition of organizations including, “publishers, resource partners, associations and other organizations involved in scholarly publishing” (“CHORUS

Update: The Association of American Publishers” 2013). CHORUS promises to provide a framework that leverages existing infrastructure to fulfill agency policies that emerge from the OSTP guidelines. The Association of Research Libraries, a major sponsor of SPARC, along with the Association of American Universities (AAU), and the Association of Public and Land-grant Universities (APLU), have drafted a separate proposal titled SHared Access Research Ecosystem (SHARE). Again, there are key differences in the proposals with implications for access to data produced from federally funded research. The advantage of the CHORUS proposal is that it would lead searchers to the final published version of the article located at the publisher’s site, while SHARE would lead to various repositories which may only contain the final edited document prior to publication. Those interested in Open Access will surely continue to debate the merits of each proposal for access to peer-reviewed publications; however, differences in how each proposal addresses data could have significant implications on data sharing in the future. Currently CHORUS does not address how the system would lead to access to the digital copies of underpinning data. Also the CHORUS system does not expose its data to external search engines, requiring users navigate to the CHORUS web site to search for articles. The SHARE proposal also seems to primarily focus on publication retrieval, but would also use the same framework to link the access points to the underpinning data. SHARE would also expose their data to external search engines providing additional points of discovery. To date it is unclear if either proposal will be adopted and how they may relate to the February 22, 2013 OSTP directive. One thing is clear, neither proposal specifies where data might be stored, backed up, and secured. The OSTP policy does not include funding for the creation of data repositories and calls for agencies to utilize infrastructure, wherever that may be.

The Role of Government

While the Obama administration and federal government are influential players, they play a different role than that of an interest group. It is the role of the government to operate as an agent of the State and to enact policies that benefit the best interest of the polis. Statements by the Obama administration indicate an interest in OA issues with specific concerns for access to data. In an Executive Order issued on May 09, 2013 President Obama states:

“To promote continued job growth, Government efficiency, and the social good that can be gained from opening Government data to the public, the default state of new and modernized Government information resources shall be open and machine readable. Government information shall be managed as an asset throughout its life cycle to promote interoperability and openness, and, wherever possible and legally permissible, to ensure that data are released to the public in ways that make the data easy to find, accessible, and usable” (“Executive Order -- Making Open and Machine Readable the New Default for Government Information | The White House” 2013).

Sections 2 and 3 of the Executive Order speak directly to requirements of federal funding agency for supporting and implementing an Open Data Policy (“Executive Order -- Making Open and Machine Readable the New Default for Government Information | The White House” 2013). On July 16, 2013 the OSTP released a sneak peak of Next.Data.gov, a site that will index datasets from agencies that publish data catalogs available for download (“First Look at Next.Data.gov | The White House” 2013).

Interest by the federal government for access to the results of scientific research, in combination with policy mandates, have proven in the past to speed up adoption of OA principles and lead to innovation in library services. The NIH Public access policy spawned new roles and services for libraries. Carpenter et al state, “Libraries used this opportunity to educate faculty, researchers, and administration about the larg-

er open access (OA) movement and to assist with compliance issues” (Carpenter et al. 2011). Similarly, a January 18, 2011 National Science Foundation policy requiring a data management plan led many libraries in STEM fields to develop service models to support researchers. As curators of knowledge sources, federal policy changes involving access to data could operate as a catalyst for new services for libraries and redefine relationships among researchers and eScience librarians (Hagenmaier, Rolando, and Rascoe 2013).

Implications for Libraries

As a major funder of research in higher education, the federal government plays a role in developing policies around Open Data that are likely to drive change in research practices at higher-education institutions. The 2011 National Science Foundation (NSF) data management requirement acted as a catalyst for new support roles for librarians. The February 2013 OSTP mandate and subsequent policies from the 11 major federal funding agencies will create even greater opportunities for library services around data support. Investigators are going to need assistance navigating the new requirements and libraries are uniquely positioned to provide services. In fact, new policies could inspire a number of new services distributed throughout the research lifecycle. Data should be thought of and curated like any other information source, whether it is a journal, book, or data set. Investigators will need services to help create data management plans for more competitive grant applications, assistance on where and how to store their data, descriptive cataloging, discoverability systems for data. According to Antell et al., “...many of the data management requirements involve the kind of work in which librarians already have expertise – organizing information, applying metadata standards, and providing access to information. For these reasons, science librarians in particular are mobilizing to meet the needs of researchers faced with the chal-

lenge of developing data management plans" (Antell et al. 2013). Instruction and bibliographic services might include new services on the use of data DOIs, data indexing systems, data citation standards, ethical considerations for the reuse data, intellectual property, and patent concerns. Repositories might be discipline specific, but could reside in institutional repositories, an area where many libraries are already providing services. The impact on access to data will filter down in to the educational mission of higher education where subject libraries will provide instructional services to undergraduate and graduate education. Access to clinical data will generate new questions that need to be navigated and has the potential to spur HIPAA reform (Goben and Salo 2013). Now is the time for libraries and librarians to retool their skill sets to accommodate emerging and evolving needs.

As with the OA movement, it is hard to predict the eventual impact of open data on library services. What is apparent is that the political economy of OA will shape research enterprises in higher education since compliance with federal policy is mandatory. At this point, the interest groups are focusing most of their efforts on peer-reviewed publications. Despite the fact that access to data has just entered the conversation on the political landscape, it is important that libraries and eScience librarians make themselves aware of the positions of interest groups since they might influence the process. Libraries have carved out new positions that specifically focus their attention on data services; however, this is not enough. Decoupling data from the publications will inspire a new generation of research that challenges the paradigm peer-review publishing model that has existed for hundreds of years. New publishing models have already emerged and concepts like Altmetrics seek to redefine how we assess impact. All academic librarians, not just eScience librarians, could be impacted as access to data will present familiar challenges in new light. Collections, liaisonship, public service, cataloging, and

instructional services will need to evolve alongside of the research process if libraries are to remain relevant in the future.

References

Antell, Karen, Jody Bales Foote, Jaymie Turner, and Brian Shults. "Dealing with Data: Science Librarians' Participation in Data Management at Association of Research Libraries Institutions." *College & Research Libraries* (Pre-Print 2013): Forthcoming July 2014. <http://crl.acrl.org/content/early/2013/04/05/crl13-464.full.pdf+html>.

Bosch, Stephen, Kittie Henderson, & Heather Klusendorf. "Periodicals Price Survey 2011 | Under Pressure, Times Are Changing." *Library Journal* 136 (8): 30–34.

Cooper, Bruce, James Cibulka, and Lance Fusarelli, eds. *Handbook of education politics and policy*. Routledge, 2008.

The Association of American Publishers. "Open Access." Accessed April 18, 2013. <http://www.publishers.org/issues/5/8/>.

The Association of American Publishers. "Publishers Introduce CHORUS." Accessed August 10, 2013. <http://publishers.org/press/110/>.

Blumer, Herbert. *Principles of sociology*. Edited by Alfred McClung Lee. New York: Barnes & Noble, 1969.

Carpenter, Maria, Jolie Graybill, Jerome O'ford Jr, and Mary Piorun. "Envisioning the library's role in scholarly communication in the year 2025." *portal: Libraries and the Academy* 11, no. 2 (2011): 659-681, <http://dx.doi.org/10.1353/pla.2011.0022>

Dyer, Owen. "Publishers hire PR heavy-weight to defend themselves against open access." *BMJ* 334, no. 7587 (2007): 227, <http://dx.doi.org/10.1136/bmj.39112.439051.DB>

The Endocrine Society. "NSF Broadens Biosketch Publications Category to Include Products; Society Recommends Similar Change to NIH." Accessed April 19, 2013. http://www.endo-society.org/advocacy/insider/2013/NSF_Broadens_Biosketch_Publications_Category_Include_Products_Society_Recommends_Similar-Change-to-NIH_Art1.cfm.

Goben, Abigail, and Dorothea Salo. "Federal research Data requirements set to change." *College & Research Libraries News* 74, no. 8 (2013): 421-425.

Hagenmaier, Wendy, Lizzy Rolando, Fred Rascoe, Dan Cohen, Christine George, Kari Watkins, Peter Suber, and Michael Chang. *Information Now: Open Access and the Public Good*. Podcast audio. October 21, 2013. <http://hdl.handle.net/1853/49236>

Hey, Tansley, Stewart, and Kristin Michele Tolle, eds. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, 2009.

Holdren, John P. "MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES." Accessed July 12, 2013. http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

International Association of Scientific, Technical & Medical Publishers. "An Overview of Scientific, Technical, and Medical Publishing and the Value it adds to Research Outputs." Accessed April 19, 2013. http://www.stm-as-soc.org/2008_04_01_Overview_of_STM_Publishing_Value_to_Research.pdf.

Joseph, Heather. "New Public Access Legislation Introduced in U.S. House of Representatives." SPARC. Accessed July 8, 2013. <http://www.sparc.arl.org/blog/new-public-access-legislation-introduced-us-house-representatives>.

Matthews, Susan. "NIH will withhold grant money to enforce public-access policy." *Nature Medicine* 19, no. 3 (2013): 3, <http://dx.doi.org/10.1038/nm0113-3>

Pusser, Brian, Ken Kempner, and Simon Marginson, and Imanol Ordorika, eds. *Universities and the Public Sphere: Knowledge Creation and State Building in the Era of Globalization*. New York: Routledge, 2012.

SPARC. "Federal Research Public Access Act." Accessed April 20, 2013. <http://www.sparc.arl.org/advocacy/national/frpaa>.

SPARC. "Why Open Access?" Accessed April 18, 2013. <http://www.sparc.arl.org/openaccess/why-oa.shtml>.

Sporkin, Andi. "AAP STATEMENT ON FASTR ACT." *The Association of American Publishers*. February 14, 2013. <http://www.publishers.org/press/94/>.

Stone, Deborah A. *Policy Paradox: The Art of Political Decision Making*. New York: W.W. Norton & Company, 2001.

Suber, Peter. "Open-Access Timeline (formerly: FOS Timeline)." Accessed April 30, 2013. <http://legacy.earlham.edu/~peters/fos/timeline.htm>.

The White House. "Executive Order -- Making Open and Machine Readable the New Default for Government Information." Accessed August 16, 2013. <http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government>.

The White House. "First Look at Next.Data.gov." Accessed August 23, 2013. <http://www.whitehouse.gov/blog/2013/07/15/first-look-nextdatagov>.

Disclosure: The author reports no conflicts of interest.

All content in Journal of eScience Librarianship, unless otherwise noted, is licensed under a Creative Commons Attribution-Noncommercial-Share Alike License <http://creativecommons.org/licenses/by-nc-sa/3.0/>

ISSN 2161-3974