

This is an Author's Accepted Manuscript of an article whose final and definitive form, the Version of Record, has been published in *The Serials Librarian*, 68: 79-86, 2015 [copyright Taylor & Francis], available online at: <http://www.tandfonline.com/doi/abs/10.1080/0361526X.2015.1023690>

OA in the Library Collection: The Challenges of Identifying and Maintaining Open Access Resources

CHRIS BULLOCK and NATHAN HOSBURGH

Presenters

SANJEET MANN

Recorder

While librarians, researchers, and the general public have embraced the concept of Open Access (OA), librarians still have a difficult time managing OA resources. To find out why, Bullock and Hosburgh surveyed librarians about their experiences managing OA resources and the strengths and weaknesses of management systems. At this session, they shared survey results, reflected on OA workflows at their own libraries, and updated audience members on relevant standards and initiatives. Survey respondents reported challenges related to hybrid OA, inaccurate metadata, and inconsistent communication along the serials supply chain. Recommended solutions included the creation of consistent, centralized article-level metadata and the development of OA collection development principles for libraries.

KEYWORDS Open Access, metadata, knowledgebases, link resolvers, discovery services, electronic resources

Managing Open Access (OA) resources remains challenging for libraries, even though librarians, researchers, and the general public have embraced the concept of OA. To find out why, Chris Bullock of Southern Illinois University Edwardsville and Nathan Hosburgh of Rollins College surveyed librarians about the systems and procedures they used to manage OA resources and how they could be improved. At their session, Bullock and Hosburgh reviewed key standards and initiatives related to OA management, reflected on the OA management practices used by survey respondents and at their own libraries, and considered how information professionals could collaborate to solve persistent problems identified in their survey. Hosburgh opened the session with a review of the Open Access environment, listing the types of OA resources and systems covered in the study. He also covered several initiatives undertaken by publishers, researchers, vendors, and librarians, demonstrating that OA drives much innovation in scholarly communication. Next he reviewed the various systems Southern Illinois University Edwardsville and Rollins College use to manage electronic resources, including OA content. Discovery services, catalogs, knowledgebases, library web pages, and research guides each come with their

own challenges for library staff, which OA resources often exacerbate. Hosburgh updated the audience on Scholarly Publishing and Academic Resources Coalition (SPARC) and National Information Standards Organization (NISO) projects intended to fully describe the spectrum of OA, and improve the metadata supply chain that undergirds OA resource management. Bullock moved on to describe the survey design and presented results in detail. Finally, both researchers offered conclusions and opened the floor for audience discussion.

SURVEY DESIGN AND RESULTS

Bullock and Hosburgh developed a 25-question survey to determine how librarians currently manage OA resources and how these processes could be improved. They asked librarians to share OA management problems, evaluate the strengths and weaknesses of management tools, and consider the value of OA resources to their overall library collections. Bullock and Hosburgh sent their survey to six electronic discussion lists frequented by e-resource librarians and received 150 responses. Most respondents (just under 80%) represented libraries supporting intensive scholarship at the Master's level or greater. However, small (less than 5,000 full-time equivalent [FTE]) and large (greater than 20,000 FTE) institutions were represented about equally, each nearly 30% of the respondent pool.

Librarians deploy a panoply of systems to manage OA content, most commonly link resolvers, Online Public Access Catalogs (OPACs), database and journal lists, discovery services and research guides. Bullock and Hosburgh asked respondents to list which products they used in each category. The most commonly named electronic resource management (ERM) systems were "none" (30%), "other" (30%, primarily Innovative Interface's ERM module), and Proquest 360 Resource Manager (23%). The most commonly named link resolvers were Proquest 360 Link (44%), Ex Libris SFX (26%), and EBSCO LinkSource (14%). The most commonly named discovery services were "none" (29%), Proquest Summon (25%), EBSCO Discovery Service (18%), and "other" (18%). Not surprisingly, respondents were vocal about the challenges of managing OA content in this multifaceted environment. Several themes emerging from the survey data were reinforced by anecdotal experiences from the session presenters and audience members.

HYBRID OA IS HIGHLY PROBLEMATIC

Many scholarly communication stakeholders are experimenting with new ways of opening access to books, articles and data sets. Author fees are only one of many revenue streams that could allow OA content to coexist alongside subscription content in an economically viable manner. Yet as one respondent bemoaned, from a management standpoint, "Hybrid OA is a nightmare!" Hybrid resources mix closed and open items within the same information containers (e.g., a journal issue), so that OA occurs at the item level rather than the title level; however, most of the tools librarians use to manage OA resources and make them discoverable, such as knowledgebases and OPACs, operate only at the title level. With no widely adopted metadata

standard or universal visual convention for denoting OA at the item level, machines and humans alike have a difficult time distinguishing open and closed items within hybrid resources.

As a testament to the difficulty of managing hybrid journals, only 10% of respondents stated their libraries usually provided access to this type of resource, twice as many said their libraries never provide access, and over a quarter of respondents were not sure how to answer.

Respondents considered discovery services to be the most effective tools at managing hybrid content, since these systems are designed to return item-level result sets. Even so, discovery services with facets designed to identify OA journal titles run the risk of excluding hybrid journals containing individual OA articles. Knowledgebases were judged to be the least helpful tools, and respondents also criticized other tools dependent on a knowledgebase—ERM systems and link resolvers—for handling hybrid content poorly. When asked what would make it easier to manage OA content in general, many respondents seized on hybrid OA, asking publishers to create and distribute article level metadata for all items describing open or closed status.

POOR COMMUNICATION

Hosburgh observed that the National Information Standards Organization (NISO) Open Access Metadata and Indicators (OAMI) working group published a draft recommendation in early 2014 addressing this very issue.¹ While the working group did not attempt a definition of “Open Access” or create a universal OA logo, it did propose two metadata elements which would identify an item as “free to read” and link to a license agreement governing use permissions for that item. The tag would indicate that an item could be viewed without payment or authentication, and would contain start and end dates to accommodate publisher embargoes. The tag would contain a Uniform Resource Locator (URL) pointing to a human or machine-readable license and indicate the start date when the license took effect. The working group recommended that all publishers create these elements during their standard editorial workflows. The elements would be incorporated into content feeds to third parties including aggregators, as well as metadata feeds to knowledgebases, linking services like CrossRef, and alerting services like electronic tables of contents (e-ToCs).

Hosburgh depicted these information flows as one half of a bidirectional, continuous process. While content and metadata flow “downstream” from publishers to aggregators to libraries and finally to their users, feedback and strategic requirements need to flow “upstream” from scholars to libraries, and ultimately to vendors and publishers.

Judging from the reactions of survey respondents and audience members, this stream is in dire need of cleanup. Respondents reported that inconsistent metadata was one of the biggest OA management challenges they faced, and requested stricter standards and better metadata for discovery when asked what would make OA management easier. During the question and answer (Q&A) session, an audience member shared an anecdote about a university press that created article-level metadata on OA and shared it with CrossRef, only to have an aggregator decline to incorporate the metadata into its indexing. She encouraged librarians not to subscribe to databases that lack necessary functionality. Another audience member questioned whether

libraries truly had the ability to influence the creation and distribution of metadata, while expressing concerns that the OAMI recommendations would not go far enough toward making OA content discoverable.

Confusion surrounding the OA environment extends beyond metadata to the very look and feel of OA content. Early in the session, Hosburgh explained how inconsistent language and visual symbols make it difficult for library users to identify OA content. The orange “open lock” icon created by Public Library of Science (PLOS) and popularized by SPARC through Open Access Week is approaching the recognizability of a web convention. However, publishers are not required to use this logo, and have devised myriad visual indicators and phrases to designate open content, including “increased access,” “public access,” or simply, “free.” In his 2013 study, Chad Hutchens documented platforms attempting to communicate OA status at the item level through Hyper Text Markup Language (HTML) <meta> tags, Dublin Core elements and even Cascading Style Sheets (CSS) formatting.²

Licensing is another problematic area for OA resources. Many library users expect OA resources to include liberal use and reuse rights for things like classroom use and data harvesting. However, publishers do not communicate these rights in a consistent manner. The OAMI <free_to_read> tag does not provide guidance about reuse rights, and the <license_ref> tag does not specify any mechanism for enforcing the license. The inconsistency can leave readers confused about what they are allowed to do with OA items. Hosburgh offered the example of one spectacular communication breakdown involving a chemistry professor who vented his frustration about a publisher’s Gold OA program in a blog post. The chemist found an interesting article but objected to the publisher’s choice of Creative Commons license (Creative Commons Attribution + Noncommercial [CC-BY-NC]), which prohibited use “for commercial purposes.” The chemist reasoned that this license barred him from assigning the article to his students, since his course enrollment generated revenue for the university that employed him. He also noted that the publisher’s “Get Permissions” link led to a generic Copyright Clearance Center request form rather than to personalized assistance. Taken together, the publisher’s idiosyncratic purple and gold OA icons, restrictive license and anonymous form comprised a “trail of difficulties” in the eyes of the professor, who concluded that the publisher lacked serious commitment to Open Access.³

Bullock and Hosburgh’s survey and discussion demonstrates that librarians long for consistent communication throughout the serials supply chain. Differing definitions and levels of commitment to OA may account for the plethora of metadata strategies, license agreements, and visual and verbal indicators currently being used to communicate OA content and permissions. While Hosburgh noted that OAMI chose not to define OA as part of its recommendations, other groups such as the Budapest Open Access Initiative (BOAI) have created influential definitions and are each pursuing work on OA management problems. For example, one audience member said that the Digital Public Library of America (DPLA) has constituted a working group to figure out rights metadata, and urged session attendees to piece together the various conversations to “get to the promised land” of interoperability and consistency.

INACCURATE INFORMATION

Another common theme running through Bullock and Hosburgh's survey involved frustration with inaccurate information in OA management systems affecting patron discovery and staff workflows.

Hosburgh noted that databases and discovery services often include facets designed to limit result sets to OA items. Librarians praised knowledgebases for helping them provide more OA resources, and appreciated link resolvers that clearly labeled OA items as "FREE" on the result screen. They were dissatisfied with interfaces whose facets missed OA items in hybrid journals, allowed irrelevant items to enter the result set, or lacked an OA filter altogether.

A knowledgebase's greatest strength—centrally managed collections—becomes a weakness when titles or holdings fall out of date and library staff have to initiate troubleshooting. Some knowledgebases do not centrally manage their OA collections, leaving it up to librarians to make manual corrections as problems are found. This can be a costly proposition; one respondent claimed that OA problems accounted for as many as 15% of all e-resource errors at their library. Other respondents maintained that OA content was no more problematic than subscription-based electronic resources.

Respondents thought content publishers and knowledgebases were best positioned to improve metadata accuracy, due to their downstream effects on ERM systems, link resolvers, and discovery services. They wanted publishers to conduct peer review of metadata before releasing them into the serials supply chain, and contribute title and holdings changes to a central repository monitored by all knowledgebase vendors.

Even when systems are error-free and metadata are accurate, respondents still felt that management systems lacked functionality. ERM systems were criticized for lack of integration with databases and inadequate reporting, which made it difficult to count the number of OA resources offered by the library. Respondents alternately praised and criticized the workflows supported by their ERM systems, and overall complained of too much work for too few staff. When asked what would make OA management easier, one respondent suggested a librarian position dedicated to OA management, while another requested "Harry Potter, the Elder wand and the help of Dobby—the free elf."

LIBRARIES NEED TO CLARIFY THE ROLE OF OA IN COLLECTIONS

According to Bullock, 71% of survey respondents consider OA materials to be part of their libraries' collections. Librarians are used to thinking about the library collection as those materials that have formally entered the library's sphere of influence and can be managed by regular acquisition, cataloging and circulation workflows. OA resources pose significant challenges to this paradigm. When asked to describe the biggest challenges of OA management, respondents most commonly mentioned the sheer number of available resources (26%), various types of "unreliability" (22%) and lack of OA collection development criteria (13%). These

responses attest to librarians' struggle with assimilating OA resources into the rest of the collection.

The extent to which librarians actively promote materials to library staff and users through various formal and informal communication strategies could be taken as an indicator of the value of those materials. Bullock noted that only 48% of respondents stated that they actively promoted OA resources (42% stated "no" and 10% were not sure). The most commonly reported methods for promoting OA resources involved meetings with faculty or staff, instruction sessions, and including OA resources in subject guides.

Respondents were mixed on whether availability of OA resources affected collection development decision making. When asked how OA resources have affected collection development practices, respondents most commonly stated that they cancelled titles or avoided purchasing OA content. Open Access has not solved libraries' budget woes, and respondents criticized university administrators who assumed as much. "The overblown promise of OA solving the serials pricing crisis has not materialized," harrumphed one respondent. Yet Bullock noted that OA resources are a helpful supplement to paid collections, particularly for smaller libraries.

Respondents also noted the lack of "truly comprehensive guides for how to include Open Access resources [in a collection]. Where to find them, how to support access, and how to discern the best resources." Bullock suggested that Jill Emery and Graham Stone's Open Access Workflows for Academic Libraries (OAWAL) wiki could help fill this need.⁴

EPILOGUE: NASIG AS SOLUTION-SEEKING VENUE

Complaints aside, 89% of respondents thought the time and effort involved in providing access to OA resources was valuable to some degree. Some respondents took the philosophical view that OA management affirmed the library's traditional role "to enable patrons to access and discover resources," or reasoned that OA resources in the collection had value from being selected by librarians. Others took the contrary view that OA management did not benefit library users since "they can find them when they click on links [in Google] anyway."

Bullock and Hosburgh's study gave respondents plenty of room to air frustrations, but also encouraged them to brainstorm solutions. This blend of irritation and ingenuity extended into the Q&A session. Audience members shared their own OA management horror stories, and one member stood to introduce himself as Brian Kelly, about to present an "open source answer" to the problem of discovering hybrid OA content at the following session. The persistence of NASIG members like Bullock, Hosburgh, and Kelly at seeking solutions to the challenges of identifying and managing OA resources allowed the session to close on a ray of hope.

ORCID

Sanjeet Mann <http://orcid.org/0000-0003-4442-1053>

NOTES

1. National Information Standards Organization, “Open Access Metadata and Indicators (OAMI),” accessed June 20, 2014, <http://www.niso.org/workrooms/oami>.
2. Chad Hutchens, “Open Access Metadata: Current Practices and Proposed Solutions,” *Learned Publishing* 26, no. 3 (2013): 159–165, doi:10.1087/20130302.
3. Peter Murray-Rust, “Can I Use Wiley’s ‘Open Access’ For Teaching? NO,” Petermr’s Blog: A Scientist and the Web, March 25, 2013, accessed June 20, 2014, <https://blogs.ch.cam.ac.uk/pmr/2013/03/25/openaccess-can-i-use-wileys-open-access-for-teaching-no/>.
4. Jill Emery and Graham Stone, “Open Access Workflows for Academic Libraries (OAWAL),” accessed June 20, 2014, <http://library3.hud.ac.uk/blogs/oawal/>.

CONTRIBUTOR NOTES

Chris Bullock is Electronic Resources Librarian at Lovejoy Library, Southern Illinois University Edwardsville, Edwardsville, Illinois.

Nathan Hosburgh is Discovery & Systems Librarian at Olin Library, Rollins College, Winter Park, Florida.

Sanjeet Mann is Arts & Electronic Resources Librarian at Armacost Library, University of Redlands, Redlands, California.