Open Scholarship Initiative Proceedings, Volume 1, 2016
doi: http://dx.doi.org/10.13021/G8FK5D

Report from the Usage Dimensions of Open Workgroup

Eric Archambault, Lorcan Dempsey, Chris Erdmann, Stephanie Fulton, Dee Magnoni, Emily McElroy, Amy Nurnberger, Ginger Strader, Michael Van Woert

Abstract

The Usage Dimensions of Open workgroup came together and considered definitions and priorities around its topic. From priorities, themes were identified. One theme included the character of research outputs and the actual research workflow process. The second theme represented economic considerations. Stakeholders were identified, and solutions considered. Solutions included both short- and long-term actions.

OSI2016 Workgroup Question

What are the usage-related challenges currently faced by open efforts? For instance, open data is intriguing in principle, but in reality, making underlying data open can be problematic, conflicting with the need for research secrecy (whether driven by the desire to be first to publish, or the desire of funders to hold onto data to protect future discovery potential), the potential for misinterpretation by other researchers, and so on. Publishing clinical trial data in open formats is also intriguing but would run afoul of many current consent agreements, particularly older consents. Open access is similarly challenged in some instances by a conflict between which versions of papers is allowed to appear in open repositories. What is the value of archiving non-final versions? What is the range of issues here, what are the perspectives, and what might be some possible solutions?

The Usage Dimensions of Open (UDO) workgroup came together in a conference room tucked down a secluded hallway on the George Mason University campus. Our lead facilitator, Amy Nurnberger, led workgroup members through an ice-breaker exercise where each member illustrated their path to the Open Scholarship Initiative conference. Members represented a diversity of backgrounds, geographic locations, and industries, and rolled up their collective sleeves to engage in two days of discussions.

The workgroup question was a broad one, and three subgroups were formed to brainstorm definitions and priorities. Subgroups then reported on top priorities, and themes were sought amongst the ideas. Two threads created a common theme: the character of research outputs and the actual research workflow process. One represented content, the other the tools, systems, and infrastructure for moving that content from idea to realization. A second theme revolved around economics. No single business model would work for all publishers, institutions, funders, governments and the like. The need

for flexible, workable economics is essential for future scholarly communications success. Even "open" as a business model has several facets, with more evolving each year.

With priorities thus defined, workgroup members turned toward stakeholder identification. The number of potential stakeholders was quite large when the breadth of interested parties was considered—from career stage, to industries, and to the press, funders, legislators, governments, and so on. Members agreed to categorize stakeholders as either primary, secondary, tertiary, or redundant. The workgroup focused on primary stakeholders. Their roles involved direct interaction in the scholarly life cycle, and changes had impact on workflow, rewards, and pain points. Secondary stakeholders were involved in the scholarly life cycle, but did not experience the impact of changes that the primary group did. Tertiary stakeholders were only involved on the periphery, and could be set aside. Redundant stakeholders had needs, barriers and challenges similar to already-defined stakeholders, and could be eliminated.

The list of eleven initial stakeholders was narrowed into four categories: researchers and librarians; funders; service providers (publishers, database and tool providers); and the public. Researchers and librarians had been two separate categories, but members agreed that in the context of open access (OA) and our defined priorities, both groups shared concerns and goals. For example, librarians and researchers are both concerned about the pain points involved in the scholarly workflow, and both want to improve this system. Both also want to maximize institutional research impact, and making publications more broadly available can provide one solution. The public as stakeholder has several variations, from citizen scientist to student, from reporter to elected official, and more. In the end, each of these stakeholders comprise net consumers of scholarly information, and all want the maximum amount of content available as soon as possible.

Understanding priorities and identifying stakeholders led to a discussion of areas in need of solutions. Subgroups were once again formed to brainstorm, and the reporting out discussion identified four areas where we could begin to look forward. First, industry standards and norms are needed across the scholarly life cycle. Second, project exit strategies are needed to provide project sustainability while accommodating the evolution of funding and ownership. Incentive systems are needed to ensure the win-win solution amongst stakeholders. Finally, the move toward open involves a complex system where inertia must be counterbalanced by a desire for change. Continuity in services and outputs must be optimized throughout the system.

With priorities, stakeholders and areas in need of solutions each defined, what actual solutions could be considered? Day two brought a lively discussion of possibilities. Could a group of major funders be brought together or could a single funder facilitate a conversation around budgeting to support OA? If funders supported OA costs within grants and awards, what would be the impact on funds for direct research? Institutions and researchers would likely have mixed feelings about a rebalancing of support. Additionally, the current promotion and tenure process in academia frequently rewards publication in high impact journals. Would added

support for OA by funders find resistance in this process? Could partnerships between funders and others address this rebalancing? An external partner such as Force11 could potentially take up the creation of an issue brief concerning funder support of OA. OSI should identify conversations that are already happening in this area, looking for synergies and potential partnerships, and facilitate knowledge sharing in this area.

Partnerships beyond funders were also considered. Intelligent collaborations among stakeholders hold great promise. No single stakeholder will accomplish OA goals in isolation. An example of coordinated, collective action could be the creation of a suite of smart tools in a machine-learning environment. Tools in the suite would learn a researcher's core funding sources, collaborators, database preferences, and workflow practices. Requirements and mandates could be met and automated in the workflow, allowing the researcher to focus on their actual research rather than filling out myriad forms, saving to relevant repositories, and reporting out results, all of which currently consume a great deal of time. Workflow would essentially become transparent. Example of tools in the suite could be Zotero, Dropbox, ORCiD, Crossref, GitHub, Zenodo, and OpenVIVO. Parallel to this suite of tools, an institution-based set of preservation tools would ensure long-term access to cited resources, research results, data sets, and other outputs of the scholarly life cycle. Citations that include hiberlinks[1] or persistent URLs that enable human and machine accessibility should become a standard practice as more smart preservation tools explored and created.[2]

A more immediate solution is to perform a landscape assessment of scholarly communications and workflow tools to categorize current best practices, standards and norms. This project may be considered by a workgroup member's organization. A good source for current practices is "101 Innovations in Scholarly Communication."[3] Project leads Jeroen Bosman and Bianca Kramer, of the Utrecht University Library, are exploring what factors drive innovation, and how workflow changes can lead to better and more open science.[4]

The workgroup defined "exit strategy" as a means of leaving one's current situation with a path to sustainability through considering appropriate governance. An example of a project that may have been sustained through a strong exit strategy is Trove, an online archive that will stop adding to its collection due to funding.[5] To address an exit strategy solution, members proposed an idea workshop to develop project solicitations. Available projects and resources such as JISC, JSTOR and Ithaka should be considered and included.[6]

Moving forward, OSI should include industry stakeholders such as Google, Microsoft and Amazon in future conferences. These stakeholders drive direct and indirect tools and services, and will strengthen the OSI conversation as well as provide partnering opportunities with other service providers, research institutions and not-for-profits.

Across stakeholder groups the need for open scholarly communications education was recognized. The workgroup identified academia as a valuable location and period to target people who would be or become

one or more type of stakeholder. Including open scholarly communications in disciplinary curricula would establish a base of understanding and a mindset that could carry into individual careers and industry paths. An obvious challenge to this solution is the already crowded curriculum of undergraduate and graduate programs. The tenure process holds its own challenges, and in its current state largely promotes the status quo in areas of scholarly communications rather than change. There is a plethora of accrediting bodies, both discipline-based and regional, which may or may not have the interest and the will to implement curriculum changes. To begin the discussion, the following groups could be invited to the table: the Association of American Universities (AAU), the Association of Public & Land-Grant Universities (APLU), and the Association of American Medical Colleges (AAMC), amongst others. 7 The discussion would include shared thoughts, concerns, and opportunities for change. Grants for curriculum and training opportunities, such as the National Institutes for Health's Broadening Experiences in Scientific Training (NIH BEST) could be sought.[8]

When exploring solutions the workgroup did not want to overlook work in progress. The National Science Communication Institute (nSCI) has a number of strong initiatives underway.[9] This work should definitely continue, with partnerships considered where relevant.

The Usage Dimensions workgroup took a tools, workflow and funding approach to our question, identifying key priorities and stakeholders, then areas that need solutions. Members worked together to craft initial solutions, identifying short- and long-term projects and goals. The two plus days of intensive meetings provided an excellent forum to come together, understand a problem from several viewpoints, and work together toward solutions. The group collectively thanks OSI for the opportunity to help craft the future of OA and scholarly communications.

OSI Usage Dimensions of Open Workgroup

Eric Archambault, President and CEO, Science-Metrix & 1science

Lorcan Dempsey, VP Research, Online Computer Library Center (OCLC)

Chris Erdmann, Head Librarian, Harvard-Smithsonian Center for Astrophysics, John G. Wolbach Library

Stephanie Fulton, Executive Director, The University of Texas MD Anderson Cancer Center

Dee Magnoni, Research Library Director, Los Alamos National Laboratory

Emily McElroy, Director, McGoogan Library of Medicine, University of Nebraska Medical Center

Amy Nurnberger, Research Data Manager, Columbia University

Ginger Strader, Director, Smithsonian Institution Scholarly Press

Michael Van Woert, Executive Officer and Director, National Science Board Office, National Science Foundation (NSF)

Notes:

[1] Hyberlnk.org, as of June 8, 2016: http://hiberlink.org

[2] Starr J, et al. (2015) Achieving human and machine accessibility of cited data in scholarly publications. PeerJ Computer Science 1:e1 https://doi.org/10.7717/peerj-cs.1

[3] Bosman, Jeroen and Bianca Kramer, "101 Innovations in Scholarly Communications," (project website), as of June 8, 2016: https://innoscholcomm.silk.co/

[4] Kramer, Bianca and Jeroen Bosman "101 Innovations in Scholarly Communications: 101 Innovations in Scholarly Communication—the Changing Research Workflow," 2015, https://dx.doi.org/10.6084/m9.figshare.1286826

[5] [ADD LINK to support] Trove, National Library of Australia, as of June 8, 2016: http://trove.nla.gov.au

[6] JISC, as of June 8, 2016: https://www.jisc.ac.uk/; JSTOR, as of June 8, 2016: http://www.jstor.org/; Ithaka, as of June 8, 2016: http://www.ithaka.org/.

[7] Association of American Universities (AAU), as of June 8, 2016: http://www.aau.edu/; the Association of Public & Land-Grant Universities (APLU), as of June 8, 2016: http://www.aplu.org/; and the Association of American Medical Colleges (AAMC), as of June 8, 2016: https://www.aamc.org/.

[8] National Institutes for Health, Broadening Experiences in Scientific Training (NIH BEST), as of June 8, 2016: http://www.nihbest.org/.

[9] National Science Communication Institute, as of June 8, 2016: http://nationalscience.org/.