Chapter 12

The Role of Repositories in the Future of the Journal

Sarah L. Shreeves

Abstract: The UK Working Group on Expanding Access to Published Research Findings, chaired by Dame Janet Finch, report, "Accessibility, sustainability, excellence: how to expand access to research publications," helped to crystallize a long simmering debate within the open access (OA) community: should the focus for OA advocates be "green" open access – that is, the use of repositories to make research published through traditional subscription-based venues openly available – or should it be 'gold' open access – that is, through publication within venues that are themselves open access? This chapter argues that that this has never truly been an either/or proposition, and that this debate often ignores or minimizes the wide variety of roles – direct and indirect – that repositories play within the larger scholarly publishing ecosystem. Research data, funder and institutional mandates for open access to published research via repositories, and the growing role of library as publisher, are all evidence that the repository – whether institutional or disciplinary or format driven – will continue to play a role within the larger scholarly publishing environment.

Keywords: Open access, institutional repository, disciplinary repository, research data management, library publishing, funder mandates

In June 2012, the Working Group on Expanding Access to Published Research Findings, chaired by Dame Janet Finch, published its report, *Accessibility, Sustainability, Excellence: How to expand access to research publications*, which recommended, among other things, that the United Kingdom move towards a set of policies that would require open access to government funded research results, specifically through publication in open access journals, so-called 'gold open-access'. In July 2012, the UK government accepted the recommendations of the Finch group report (though inquiries into the details of implementation are ongoing at the point of writing this chapter). (UK Dept. of Business, Innovation, and Skills 2012)

Both the report and the UK government's reaction prompted a storm of commentary and consternation from various corners of academia including advocates of the use of repositories to provide open access to the published literature.[1] The reason for their consternation? While the Finch report does call for strengthening institutional and disciplinary repositories, it is clear that the Working Group sees the role of repositories as complementary to the publishing enterprise and their value primarily in providing access and preservation to research data and grey literature, rather than open access to the published literature. The Finch report notes that repositories, outside of a handful of disciplinary repositories, have been underutilized for this purpose and that 'it is unlikely that either institutional or subject-based repositories could *by themselves* provide a satisfactory model for a research communications system that involves the effective publication and dissemination of quality-assured research findings.' [emphasis in the original] (Working Group, 2012: 95)

The Finch report and the subsequent turmoil crystallized a long simmering debate within the open access (OA) community: should the focus for OA advocates be 'green' open access – that is, the use of repositories to make research published through traditional subscription-based venues openly available – or should it be 'gold' open access – that is, through publication within venues that are themselves open access? Of course, this has never truly been an either/or proposition (though one might be forgiven for not knowing that given the rhetoric by some in the OA community), and this debate also often ignores or minimizes the wide variety of roles – direct and indirect – that repositories play within the larger scholarly publishing ecosystem.

Readers of the first edition of *The Future of the Academic Journal* will know that I titled my chapter 'Cannot Predict Now' – a reference to the answer given by Mattel's Magic Eight Ball when asked a yes or no question. (Shreeves, 2009) I venture now, some four years later, to give a more positive prediction, though perhaps not in the ways that I had initially envisioned. Data publication was (literally) just a footnote in the last edition. Institutional mandates for open access (requiring the use of institutional repositories) had just been established at Harvard's Faculty of Arts and Sciences, and had not yet spread to over 150 institutions worldwide (as recorded by ROARMAP at *http://roarmap.eprints.org/*). Repositories were primarily institutional or disciplinary. All of this has changed dramatically and substantively.

As I did in the last edition, I make the disclaimer that I manage an institutional repository and thus have certain prejudices and hopes. I have tried to minimize these in the presentation of this discussion.

The current repository landscape

As described in the first edition, a repository is, at its core, a 'digital assets management system of some kind or a network of systems that allows for the deposit and subsequent distribution of digital files over the internet.' (Shreeves, 2009: 198) But beyond this very basic description, repositories vary wildly in terms of content, access, management, communities, and purpose; though, for the purposes of this discussion, I will be limiting my description of the landscape to those that have some relationship to scholarly publishing and communications.

Conventional repositories have focused on content, often in the form of scholarly papers, generated by researchers either within a specific institution or set of institutions (*institutional repositories*) or for specific disciplines (*disciplinary repositories*). The OpenDOAR directory (*http://www.opendoar.org/*) lists over 2,200 institutional and disciplinary repositories. But the landscape is much more diverse than even four years ago. Repositories focused on content such as those for data, code, even scientific protocols are now more common; these are sometimes quite closely linked to publishers. There are also services such as Mendeley and Academia.edu, whose primary focus may not be their repository role but who do allow researchers to upload their research and share it (often just to specific groups). These changes have made it more difficult to define and count what is a repository. While OpenDOAR and the Registry of Open Access Registries (ROAR; see *http://roar.eprints.org/*) can both be used to explore numbers and characteristics of conventional repositories worldwide, they do not list all repositories, particularly those that may not explicitly describe themselves as such.

Content within repositories can include: some version of published journal articles, book chapters, and books; conference papers and posters; theses and dissertations; video and audio files of talks and lectures; original undergraduate research; technical reports, white papers, and other forms of grey literature; research data in all formats; and code, scripts, and software. The repository, IDEALS (*http://www.ideals.illinois.edu/*), for which I am responsible at the University of Illinois at Urbana-Champaign, contains examples of all of these. For many institutional and disciplinary repositories, the major focus continues to be document like objects whether those are grey literature, theses and dissertations, or the published literature. There is generally no peer or editorial review of the material deposited in the repository; for an institutional repository, the author's affiliation combined with a repository's collection policy is usually sufficient for accepting a deposit. Disciplinary repositories have a variety of methods to vet deposits; arXiv, for example, may require that a potential depositor be endorsed by a current contributor (see *http://arxiv.org/help/endorsement*). More repositories are beginning to include means for commenting or otherwise interacting with a contribution, as well as measures of impact; for example, figshare (*http://www.figshare.com/*), a repository that allows researchers to share and make available supplemental material to an article, allows comments, 'shares' via Twitter and other social tools, and downloads.

Most repositories provide open access to the majority of the content within them; where there are restrictions, they are often in the form of limited embargos or restrictions to a specific community (for example, the campus in the case of an institutional repository). Generally, the descriptive information, or metadata, is openly available for indexing via general and specialized search engines such as Google or Google Scholar; some are also open for harvesting via protocols such as the Open Archives Initiative Protocol for Metadata Harvesting (OAI PMH) or via specific application programming

interfaces (APIs). While the content and metadata may be, for the most part, openly available, most content in repositories – particularly within institutional and disciplinary repositories – is not free from restrictions on use; in most cases, the authors or publishers still exert full copyright protections over the content. This is less true in some of the more specialized repositories, particularly those connected to the Open Science movement, defined by Peter Suber as 'combining OA texts, open data, and open-source software and providing these sorts of openness at every stage of a research project, not just at the end in reporting results' (Suber, 2012: xi); these repositories often use Creative Commons licences (*http://www.creativecommons.org/*) in order to explicitly allow reuse of code, data, and research outputs.

Repositories also have varying commitments to the long-term preservation of the content within them. In their 2011 study of 72 research libraries with institutional repositories, Li and Banach found that 97.4 per cent included preservation within the mission of the repository, though how well preservation activities are actually integrated into the repository has been questioned. (Li and Banach, 2011; Rieh et al. 2008) In some well-established disciplinary and content repositories such as arXiv (*http://arxiv.org/*) or ICPSR (*http://www.icpsr.umich.edu/*) there is an explicit commitment to perpetual, persistent access to the materials deposited, but this can be more difficult to assess in others where the primary focus is access and networking between researchers. As repositories take on roles of providing access to supplementary materials (such as data or code) to journal articles, the lack of attention and clarity concerning preservation strategies and infrastructure becomes of increasing concern.

Repositories and open access to the published literature

Most of the debate about repositories and their relationship to scholarly publishing has occurred within the sphere of the repository's role in providing open access to the published literature. Open access via self-archiving (i.e. 'green' OA) relies on repositories for the infrastructure to provide this access (though a similar result can be achieved by simply posting papers on a web site). (Harnad et al, 2004) The research is (or may be) published through traditional means, but the author(s) retain the rights to make some version of that research available online. Generally, the version that may be made available is either the preprint or the postprint. The preprint can be defined as the version prior to submission or as submitted for publication, but not yet reviewed, edited, and/or published; this would be the Author's Original or Submitted Manuscript under Review under the definitions recommended by the National Information Standards Organization (NISO). The postprint is generally either the author's final manuscript version (post peer-review but before final formatting and editing) or the published version (with the formatting, pagination, and headers and footers of the journal); this would be the Accepted Version or the Version of Record under the NISO recommendations. (NISO, 2008) As of January 2013, according to the Sherpa RoMEO directory of publisher policies on self-archiving (*http://www.sherpa.ac.uk/romeo/*) some 68 per cent of the 1,196 publishers included allow some version of an article to be self-archived; 60 per cent allow the Accepted Version to be deposited, and 25 per cent allow the Version of Record to be deposited. These numbers include both very large publishers (such as Elsevier), as well as small publishers, university presses, and scholarly societies of all sizes.

While these numbers may appear to be good news for repositories focused on open access to the published literature, a number of factors complicate the reality. Publisher's policies, while technically

Final author manuscript version of book chapter "The Role of Repositories in the Future of the Journal" to appear in The Future of the Academic Journal, 2nd edition ed by Bill Cope and Angus Phillips. Cambridge, UK: Chandos Publishing. 2014. (24 pp in typescript)

allowing self-archiving, can include any number of conditions, such as embargos and restrictions on the type of repository into which an author can deposit. Authors must understand these policies, and, in some cases, negotiate for more specific rights to make their work openly available. Authors must be comfortable, in most cases, making available a version of the research which is not the Version of Record. Authors must be motivated to undertake the work of deposit in the first place. For many disciplines, open access is not an expectation, and early sharing of research is not common or accepted outside of traditional conference venues. These factors are not trivial impediments; current estimates are that only between 11 and 15 per cent of total research output is freely self-archived. (Björk et al, 2010; Suber, 2012: 58) Disciplines in which open sharing of research before publication is common (such as high energy physics and economics) may see a much higher percentage of deposit, but these appear to be the exception to the rule. This is despite some evidence that researchers support the principles of open access to academic research. (Repository Support Project, 2011) Whatever the reason for low deposit rates – lack of direct incentives, difficulty in determining rights issues, little awareness of repositories – it seems clear that even with directed and concerted outreach most repositories are unlikely to affect the scholarly publishing system any time soon.

But what if researchers were required to make their research openly available by their institutions or funders? In early 2008, Harvard University's Faculty of Arts and Sciences (FAS) voted unanimously to do just this; the faculty gave the university rights to make their published articles available in Harvard's institutional repository, DASH (*http://dash.harvard.edu/*). For the text of the resolution see *http://osc.hul.harvard.edu/hfaspolicy*. According to the Registry of Open Access Repositories Mandatory Archiving Policies (ROARMAP: *http://roarmap.eprints.org/*), some 138 institutions and academic departments and colleges worldwide of all sizes and types have established such policies

since 2008 including Duke University, the Massachusetts Institute of Technology (MIT), Oberlin College, the University of Kansas, the National Institute of Oceanography, Queensland University of Technology, Trinity College Dublin, and the World Bank. In the majority of cases, particularly those within academic institutions, these policies have been driven by the faculty themselves and are passed by faculty senates; this is of particular import as the impetus for change is not driven by the library (although certainly the library may be the instigator), but by the producers, editors, and peer reviewers of the research.

Typically these policies (which usually grant the institution non-exclusive, joint copyright in a work) allow an author to opt out if a publisher will not allow self-archiving and do not specify where a researcher must publish. Indeed, a proposed open access resolution (not a policy, that is, non-binding) at the University of Maryland failed in part because it specifically encouraged faculty to publish in open access journals where practical. (Hackman, 2009) Institutions with open access policies provide support through provision of addendums for publishers' copyright transfer agreements, harvesting of content directly from publisher sites (where possible), and negotiation with publishers to mitigate the need for the institution's authors to negotiate to retain rights. (Duranceau and Kriegsman, 2013) MIT, for example, has negotiated with Springer to allow articles by MIT affiliated authors to be 'archived and/or deposited in any repository, or used for any scholarly or educational purposes.'[2]

Data on the overall impact of these policies is not readily available; many institutions are still in the process of implementing these. Harvard's FAS and MIT report that only 5 per cent and 4 per cent respectively of the articles have had an opt-out waiver issued. (Duranceau and Kriegsman, 2013: 92) According to a press release from the MIT News Office, roughly 33 per cent of the some 7000 articles

published by MIT faculty since 2009 have been made openly available in DSpace@MIT.[3] Publishers

have reacted in a variety of ways; MIT publicly tracks publisher cooperation with their open access

policy; most publishers listed do cooperate with the exception of some of the largest publishers, such

as Elsevier, Wiley-Blackwell, and the American Chemical Society.[5] Elsevier explicitly require authors

to opt out of any mandatory policies, despite allowing the deposit of the Accepted Version. Elsevier's

copyright policy reads that:

> …deposit in, or posting to, subject-oriented or centralized repositories (such as
>
> PubMed Central), or institutional repositories with systematic posting mandates is
>
> permitted only under specific agreements between Elsevier and the repository, agency
>
> or institution, and only consistent with the publisher's policies concerning such
>
> repositories. Voluntary posting of AAMs [Accepted Author Manuscript] in the arXiv
>
> subject repository is permitted.[4]

The reasoning behind this claim that these policies represent systematic distribution of their research,

while voluntary deposit does not (one would guess because voluntary deposit is so uneven).

Despite the deliberate lack of cooperation by some publishers, many open access advocates have

seized on the development of institutional open access policies as key to the future for 'green' OA.

However, the rate of adoption does appear to be slowing. According to ROARMAP, since a peak in

2009 where there were over 50 institutional policies passed, in 2010 the number had dropped to

approximately 30; 2011 and 2012 each saw fewer, between 20 and 30. This may be due to the

difficulty in getting such policies passed through faculty senates, interest in better understanding the impact of such policies, or because more attention has been focused on 'gold' open access.

An arguably more powerful driver of 'green' open access are funders. Funders have shown interest in open access in order to increase the impact of the funded research and, in the case of government-sponsored research, in order to allow the public access to research funded through tax dollars. In late 2005, the Wellcome Trust, the largest private funder of medical research in the United Kingdom (UK), required that its grantees make openly available published research through the UK PubMed Central. Since then 56 funders worldwide, both public and private, have instituted similar requirements for their grantees (see ROARMAP *http://roarmap.eprints.org/*) Perhaps the largest of these has been the US National Institutes of Health (NIH) which instituted a public access policy in 2008 that requires published articles be made openly available in PubMed Central a year after publication.

It is these funder mandates, particularly from government entities, that have attracted the most attention from publishers, researchers, politicians, and the public. In the United States, legislation that would expand the NIH policy to all federal agencies expending funds above a certain threshold (the Federal Research Public Access Act or FRPAA) has been introduced several times in the House and the Senate since 2006, most recently in February 2012 (H.R. 4004 and S. 2096), but has not yet seen success. New legislation that differs slightly from FRPAA was introduced into both the House and Senate in February 2013; in addition to open access to federally funded research, the Fair Access to Science and Technology Research (FASTR) Act (H.R. 708 and S. 350) would require that the

research be in a form that is easily accessible for computational analysis, and that agencies consider requiring licensing terms that would allow reuse of research with attribution.

The FASTR Act comes on the heels of a very active year in terms of open access initiatives. In November 2011, the Office of Science and Technology Policy of the White House issued two public requests for information (RFI): one on *Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research* and the other on *Public Access to Digital Data Resulting From Federally Funded Scientific Research.*[6] 378 comments were submitted for the RFI on public access to publications, and 118 on the public access to data RFI.[7] In December 2011, the Research Works Act (H.R. 3699), legislation that would roll back the NIH policy and prohibit federal agencies from enacting such policies, was introduced in the US House by Representative Darrell Issa (R-CA) and Carolyn Maloney (D-NY); it had the support of the American Association of Publishers (AAP). This legislation met with a storm of protest from researchers, libraries, societies (including the Modern Language Association and the American Physical Society), and some publishers (including MIT Press, Nature Publishing Group, and the American Association for the Advancement of Science (AAAS)). (Howard, 2012) Opponents noted that both sponsors had received campaign contributions from Elsevier. (Taylor 2012) A website, The Cost of Knowledge (*http://thecostofknowledge.com/*), was established by Sir Timothy Gowers from Cambridge University in late January 2012 to protest against Elsevier's business practices, high prices for journals, and support for acts such as the Research Works Act (as well as the Stop Online Piracy Act (SOPA) and the Protect IP Act (PIPA)); on it, academics could pledge to not publish, edit, or peer review journals published by Elsevier. This garnered immediate support by academics with over 7,000 signatories by late February. (Howard 'Legislation', 2012) By late February, Elsevier explicitly withdrew its support for the act, and the two

sponsors stated that they would not pursue it. In May 2012 a petition to the White House to require free access over the Internet to scientific journal articles arising from taxpayer-funded research was begun;[8] despite reaching the requisite 25,000 signatures in the time allotted that would guarantee a response from the White House (and despite reaching over 63,000 signatories at this writing), there has yet to be a formal response from the White House.

So where do repositories fit into all of this? According to an analysis of open access policies from 48 funders conducted by SPARC Europe, in most cases funders mandate that researchers make their research openly available via repositories rather than the 'gold' route - publishing in an open access journal or utilizing the so-called hybrid model of publishing an open access article within a subscription based journal.[9] Of all funders, it is only the Research Councils UK (RCUK), under the recommendations of the Finch Report, who would require open access via the 'gold' route and would not consider direct deposit in an open access repository sufficient (although a journal that deposits research into an OA repository would meet the requirements). Funding for Article Processing Charges (APCs), utilized by many open access journals and hybrids models, would come out of the RCUK budget, and there is an explicit expectation that institutions of higher education and research would set aside funds for APCs.[10] The dismay of many open access advocates at the Finch Report and the RCUK policy was based primarily on three factors: 1) it dismisses the infrastructure already in place for open access, a well-established network of institutional repositories, built over the last 10 years; 2) it focuses too much on assuaging publisher fears of open access via the 'green' model, and thus, presents a much more expensive model for open access; and 3) it is out of step with the majority of funder policies worldwide. (SPARC Europe 2012) The fears mentioned in the second point essentially centre on whether by providing open access to published research via repositories (even if just the

Accepted Version) subscription rates to journals would fall (via libraries cancelling) forcing publishers to shut down journals, or shift business models. There is little evidence to date to either prove or disprove this fear, although often the example of physics and the arXiv repository is used, as I did in the last edition of this book, to illustrate that active 'green' open access can exist alongside subscription based journals. (Shreeves 2009) While it is unclear whether the physics example can be generalized, the publishing industry itself provided some evidence that it is less likely that libraries will cancel journals because of availability of articles via OA repositories. In 2006 a survey commissioned by the Association of Learned and Professional Society Publishers found that librarians were more likely to consider faculty needs, usage, and price rather than open access when considering journal cancellations. (Ware 2006) With the introduction of the FASTR Act in the US which favours the use of repositories, this debate of how open access should be enacted is likely to continue for some time.

Further impact of repositories

Beyond repositories and open access, two additional areas should be mentioned when considering the impact of repositories on the future of academic journals. The first is that repositories have enabled the growth of infrastructure and expertise for publishing within libraries and other institutions. The second is the role that repositories are beginning to play vis-à-vis research data.

Infrastructure and capacity for publishing

Repositories have enabled libraries in particular to develop and expand expertise in the area of advising authors on copyright and publishing issues; for many libraries, it is expected that the liaison or subject librarians will be able to speak to these issues with their faculty. (Radom et al, 2012: 14) It would appear that the lone repository manager so aptly described by Salo is slowly disappearing. (Salo 2008) Repositories have in addition allowed libraries to develop an extensible, stable infrastructure for access, preservation, and publishing of digital content. This combination of expertise and infrastructure has meant that more libraries are experimenting with providing publishing services either out of the repository or using an aligned software product. In a 2012 survey of major research libraries, 75 per cent offered consultation support for faculty on scholarly communication issues such as retention of rights when publishing, 75 per cent offered repositories to host and preserve digital content, and 68 per cent offered 'digital scholarship' support, meaning both publishing and use of new technologies to create rich online scholarship. (Radom et al, 2012) A white paper from the National Institute for Technology in Liberal Education (NITLE) notes that the majority of its members (which are, for the most part, liberal arts institutions) have digital repositories in place. (Alexander, 2011: 21) The infrastructure to support publishing is expanding. The Hathi Trust, a partnership of over sixty academic and research libraries, which preserves and provides access to over 10 million digitized volumes is currently developing tools and infrastructure to support journal publishing directly through the Hathi Trust Digital Library.[11] BePress, in addition to providing Digital Commons (a commercial repository software) also provides software that allows libraries to support journal (and monograph) publishing; their website lists over 550 journal titles published by their customers. The Open Journal System (OJS) software package can push journal content directly into a

repository for preservation services. Libraries and librarians are also forming communities of practice around library publishing. In 2011, a The Humanities and Technology (THAT) Camp was held in Baltimore, MD that focused on libraries and publishing. Out of that daylong unconference, a listserv was formed and remains active.[12] A group blog focused on library publishing was announced in 2013.[13] Also in 2013, the Library Publishing Coalition, made up of over fifty academic libraries, was formed in order to provide a space for development of best practices, training, and research into library publishing.[14] Clearly, library publishing is expanding; while it is too soon to tell what the impact might be on the publishing system as a whole, it certainly expands the options that researchers have both in terms of outlets for publishing and starting a journal.

Data and Repositories

In January 2011 the US National Science Foundation (NSF) established a requirement that all grant proposals must include a data management plan.[15] Other funding agencies followed suit including the Office of Digital Humanities of the National Endowment for the Humanities (NEH), the Institute of Museum and Library Services (IMLS), and the National Oceanic and Atmospheric Administration (NOAA). There are similar requirements from funders in other countries as well, including the Research Councils UK.[16] Although the details of data management plans might differ among funders (and within the NSF even among the directorates and divisions), the basic elements generally include information about what data will be produced, how it will be stored and described, rights issues, and how it will be shared with other

researchers. While most plans do not explicitly require that data be made openly available or published, it is expected that data can be accessed by other researchers through some means.

While some research institutions and libraries were certainly working on research data curation and management prior to January 2011, the NSF requirement prompted much wider awareness and activity around these issues. Institutions now have to support infrastructure in order for their researchers to manage data in accordance with their plans. Well-established repositories serving specific disciplines, such as the Inter-university Consortium for Political and Social Research (ICPSR) for social science research or GenBank for DNA sequences, partially serve the publishing and access role, but for many disciplines there exists no such repository. Institutions are leveraging already existing institutional repositories, or are building out new infrastructure that better support the specific requirements of research data.

The specific issue of providing access to data through a repository or through publication is quite complicated. There are questions of what version of the data (raw? cleaned? Processed? Analysed?) should be shared. For many disciplines privacy and confidentiality issues loom large, and the effort required to anonymize and clean datasets can be resource intensive. Even when datasets are anonymized, there is the question of whether personally identifiable information could be made available through the aggregation of multiple datasets. Standard descriptions of research data is incredibly uneven, and highly dependent on whether a discipline expects datasets to be shared and published. For example, the International Union of Crystallography (IUC), requires its authors to submit a Crystallographic Information File (CIF) if describing a crystal structure.[17] Research data also comes with perhaps a more complicated set of rights issues than

traditional research and scholarship products might; although factual data is not copyrightable, most institutions in the US do claim ownership of research data developed through federal funding in a way that they do not over traditional academic work. Researchers and research institutions working with private and commercial entities may enter into contracts that prevent the sharing and publication of data.

All of these complications have meant that, in general, journal publishers have been slow to actually take on the publishing and sharing of research data sets beyond what they have traditionally done in the form of graphs, charts, and figures. A broad coalition of libraries, research institutions, and national organizations, and notably not publishers, have led the development of citation standards via DataCite (*http://www.datacite.org/*). A presentation from Elsevier for the International Association of Science, Technology, and Medical Publishers shows that the publisher is actively linking between journal articles and the referenced data sets within repositories. (Koers 2012) In January 2013, the Public Library of Science (PLoS) announced that it had partnered with figshare (*http://www.figshare.com*) to host supplemental data for all seven of its journals, including its largest, PLoS One.[18] Like the IUC mentioned above, other journal publishers within the life sciences will require the deposit of sequencing or other datasets into repositories like the Protein DataBank or GenBank. While this is still a very active area of development, the trend does appear towards the use of repositories for hosting datasets rather than the publication of datasets through academic journals directly; the expertise and resources needed for publication, display, and manipulation of datasets may be enough of an obstacle for most publishers. Whatever the reason, the growing importance of datasets as a primary output of

a research paper (that can be cited, impact measured, etc.) does indicate a point of partnership between journal publishers and repositories.

Research data, funder mandates for open access to published research via repositories, and the growing role of library as publisher, are all evidence that the repository – whether institutional or disciplinary or format driven – will continue to play a role within the larger scholarly publishing environment. As someone who has managed a repository and been active in writing and speaking about scholarly communication issues since 2005, it is fascinating to see how far we have moved from the simplistic view that the repository would change the economics of academic journal publishing; the reality is, of course, far more complicated, and, in my view, interesting. I look forward to what the next four years will bring.

Notes

1 See the listserv archives from June 2012 onwards of the JISC-Repositories listserv (*https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=JISC-REPOSITORIES*) and the Global Open Access List (GOAL) (*http://mailman.ecs.soton.ac.uk/pipermail/goal/*) for examples.

2 See *http://libraries.mit.edu/sites/scholarly/mit-open-access/open-access-at-mit/mit-springer-author-rights-agreement/.* (accessed 5 February 2013)

3 See *http://web.mit.edu/newsoffice/2012/the-worldwide-impact-of-open-access-to-mit-faculty-research.html.* (accessed 5 February 2013)

4   See *http://www.elsevier.com/about/open-access/open-access-policies/article-posting-policy#accepted-author-manuscript*. (accessed 5 February 2013) As an aside, the last sentence is rather interesting and perhaps represents an acceptance on Elsevier's part that researchers who deposit to arXiv would do it despite any prohibition otherwise.

5   See *http://libraries.mit.edu/sites/scholarly/mit-open-access/open-access-at-mit/mit-open-access-policy/publishers-and-the-mit-faculty-open-access-policy*. (accessed 5 February 2013)

6   See *https://www.federalregister.gov/articles/2011/11/04/2011-28623/request-for-information-public-access-to-peer-reviewed-scholarly-publications-resulting-from* and *https://www.federalregister.gov/articles/2011/11/04/2011-28621/request-for-information-public-access-to-digital-data-resulting-from-federally-funded-scientific*. (accessed 8 February 2013)

7   See *http://www.whitehouse.gov/blog/2012/01/30/your-comments-access-federally-funded-scientific-research-results* (accessed 8 February 2013) for the comments.

8   See *https://petitions.whitehouse.gov/petition/require-free-access-over-internet-scientific-journal-articles-arising-taxpayer-funded-research/wDX82FLQ* (accessed 8 February 2013).

9   See *http://sparceurope.org/analysis-of-funder-open-access-policies-around-the-world/* (accessed 15 February 2013).

10   See *http://www.rcuk.ac.uk/media/news/2012news/Pages/120716.aspx* for the RCUK policy. (accessed 15 February 2013)

11   See the mPach Project at *http://www.hathitrust.org/mpach* (accessed 15 February 2013)

12 See the Libpub Google Group at *https://groups.google.com/forum/#!forum/libpub* (accessed 15 February 2013)

13 See *http://librarypublishing.wordpress.com/* (accessed 15 February 2013)

14 See *http://www.educopia.org/programs/lpc* (accessed 15 February 2013)

15 See *http://www.nsf.gov/bfa/dias/policy/dmp.jsp* (accessed 15 February 2013)

16 See the Digital Curation Centre's overview of UK funder policies at *http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies* (accessed 15 February 2013)

17 For example, see the author guidelines for the *Foundations of Crystallography* at *http://journals.iucr.org/a/journalhomepage.html* (accessed 15 February 2013)

18 One should note that figshare is supported by Digital Science, a division of Macmillan Publishers Limited, although it retains its autonomy.

Bibliography

Alexander, Bryan. (2011) *The New (In)Visible College: Emergent Scholarly Communication Environment and The Liberal Arts*. Georgetown, TX: NITLE. *http://www.nitle.org/live/files/34-the-new-invisible-college* (accessed 15 February 2013)

Björk B.C., Welling P., Laakso M., Majlender P., Hedlund T., and Guðnason, G. (2010) Open Access to the Scientific Journal Literature: Situation 2009. PLoS ONE 5(6): e11273. doi:10.1371/journal.pone.0011273. Available: *http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0011273* (accessed 2 February 2013).

Duranceau, E. and Kriegsman, S. (2013) 'Implementing open access policies using institutional repositories.' In P. Bluh and C. Hepfer (eds) *The Institutional Repository: Benefits and Challenges*. Chicago, IL: American Library Association, pp. 75-97. *http://www.ala.org/alcts/sites/ala.org.alcts/files/content/resources/papers/ir_ch05_.pdf* (Accessed 3 February 2013)

Hackman, T. (2009) 'What's the opposite of a Pyrrhic victory: Lessons learned from an open access defeat.' *College and Research Library News* 70(9): 518-521, 538.

Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y, Oppenheim, C., Stamerjohanns, H., and Hilf, E. (2004) 'The access/impact problem and the green and gold roads to open access', *Serials Review* 30(4): 310–14.

Howard, J. (2012) 'Wired Campus: Some Associations, Scholars Protest Bill That Would Curb Public Access to Research' *The Chronicle of Higher Education* 25 January 2012. *http://chronicle.com/blogs/wiredcampus/some-associations-scholars-protest-bill-that-would-curb-public-access-to-research/35166* (accessed 8 Feb 2013)

Howard, J. (2012) 'Legislation to Bar Public-Access Requirement on Federal Research Is Dead' *The Chronicle of Higher Education* 27 February 2012. *http://chronicle.com/article/Legislation-to-Bar/130949/* (accessed 8 February 2012)

Koers, Hylke. (2012) *Elsevier Data and Publications Linking*. Oxford, UK: International Association of Science, Technology, and Medicine Publishers. *http://www.stm-assoc.org/2012_11_19_Research_Data_Group_Koers_Elsevier_Data_and_Publications_linking.pdf* (Accessed 15 February 2013)

Final author manuscript version of book chapter "The Role of Repositories in the Future of the Journal"
to appear in The Future of the Academic Journal, 2nd edition ed by Bill Cope and Angus Phillips.
Cambridge, UK: Chandos Publishing. 2014. (24 pp in typescript)

Li, Y. and Banach, M. (2011) 'Institutional Repositories and Digital Preservation: Assessing Current Practices at Research Libraries' *D-Lib Magazine* 17 (5/6), available at *http://www.dlib.org/dlib/may11/yuanli/05yuanli.html* (accessed 3 Feb 2013)

National Information Standards Organization (NISO). (2008) *Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group*. Baltimore, MD: National Information Standards Organization, available at http://www.niso.org/publications/rp/RP-8-2008.pdf (accessed 13 Dec 2013)

Radom, R., Feltner-Reichert, M., and Stringer-Stanback, K. (2012) *SPEC Kit 332: Organization of Scholarly Communication Services*. Washington, D.C.: Association of Research Libraries.

Repository Support Project. (2011) *Survey of academic attitudes to open access and institutional repositories: an RSP and UKCoRR initiative*. Available at *http://rspproject.files.wordpress.com/2011/12/attitudes-to-oa-basic-summary-report.doc* (Accessed 3 February 2013)

Rieh, S.Y., St. Jean, B., Yakel, E., Markey, K., and Kim, J. (2008) 'Perceptions and Experiences of Staff in the Planning and Implementation of Institutional Repositories', *Library Trends* 57(2): 168-190.

Salo, D. (2008) 'The innkeeper at the roach motel', *Library Trends* 57(2): 98-123.

Shreeves, S.L. (2009) ''Cannot predict now': The role of repositories in the future of the journal'. In B.Cope and A.Phillips (eds) *The Future of the Academic Journal*. Oxford, U.K.: Chandos Press, pp. 197-211.

SPARC Europe. (2012) A thumbs up for Open Access, but an expensive way of getting there. *http://sparceurope.org/sparc-europe-response-to-the-finch-report/.* (Accessed 15 February 2013)

Suber, P. (2012) *Open Access*. Cambridge, MA: The MIT Press.

Taylor, M. (2012) 'Academic publishers have become the enemies of science' The Guardian Online. 16 January 2012. *http://www.guardian.co.uk/science/2012/jan/16/academic-publishers-enemies-science* (accessed 8 February 2013)

United Kingdom Department of Business Innovation and Skills. (2012) Government Response to the Finch Group Report: 'Accessibility, sustainability, excellence: how to expand access to research publications', available at *http://www.bis.gov.uk/assets/biscore/science/docs/l/12-975-letter-government-response-to-finch-report-research-publications.pdf* (accessed 7 January 2013)

Ware, Mark. (2006) *Summary and Conclusions of the ALPSP Survey of Librarians on Factors in Journal Cancellations*. London: Association of Learned and Professional Society Publishers. *http://www.alpsp.org/Ebusiness/Libraries/Publication_Downloads/libraryreport-summary.sflb.ashx?download=true* (Accessed 15 February 2013)

Working Group on Expanding Access to Published Research Findings. (2012) *Accessibility, sustainability, excellence: how to expand access to research publications*. London, UK: Research Information Network, available at: *http://www.researchinfonet.org/wp-content/uploads/2012/06/Finch-Group-report-FINAL-VERSION.pdf* (accessed 7 January 2013)