Summer July 1, 2016

Heard on the Net: Open Access Rhapsody

Jill Emery

ADVISOR REPORTS FROM THE FIELD

Heard on the Net

Open Access Rhapsody

doi:10.5260/chara.18.1.41

Jill Emery *(Collection Development Librarian, Portland State University)* <jemery@pdx.edu>

Abstract

In 2012, Philip Campbell, the editor-in-chief at *Nature,* noted that Open Access to scientific research is "very compelling." In 2014, David W. Lewis wrote a compelling article for *C&RL* entitled: "The Inevitability of Open Access." For most North American librarians in the past two years, the big deals have endured and there appears to be little change in the United States. While many librarians keep an eye out for various initiatives underway and see colleagues experimenting here and there with article processing charges (APCs), these efforts are largely seen as experiments and not as new ways of doing academic scholarship in North America. In this column, three impacts are noted which take us closer to Lewis' 'inevitability' and outline how shifts are happening that could have larger ripple effects.

The Scholarly Communication Landscape & Changing Research Workflows

<https://figshare.com/articles/Views_on_innovation_The_scholarly_communication_landscape_and_changing_research_workflows_/3185293>

In April 2016, I had the privilege to attend the Force16 meeting here in Portland, Oregon. Force11 has been existence for the past five years starting in 2011 and thus, the name of the group. The folks involved with this group are a community of scholars, librarians, archivists, publishers, and research funders. They are committed to facilitate the change toward improved knowledge creation and sharing. This was a fantastic event, highlighting many extremely talented people from around the globe developing the path forward to improve knowledge creation, and change how we communicate scholarship and research. The diversity of knowledge and talent that was brought together for this meeting was truly astounding and cannot be praised and championed enough. In particular, two members, Jeroen Bosman, Scholarly Communication Specialist at Utrecht University Library and Bianca Kramer, the subject specialist, Life Sciences and Medicine at Utrecht University Library have compiled a global survey on scholarly communication that has captured the scholarly production and sharing tool usage patterns of researchers and others throughout the world. Their research is based on 20,633 survey respondents in multiple disciplines worldwide. The respondents to the survey were academic faculty members, postdocs, students, librarians, publishers, and members of industry and government as well as those identified as from the general population. The array of tools currently employed for both the production and sharing of scholarship is vast and somewhat mind-boggling. Bosman and Kramer focused specifically on the tools used for archiving and sharing publications and found that ResearchGate was by far the most used platform followed by Institutional Repositories. Three goals identified for science and scholarship were "good," meaning research that declared competing interests, could be reproduced, provided meaningful assessment, included effective quality checks, gave credit where credit was due, and did not engage in fraud or plagiarism. The second goal was that science/scholarship production was open; meaning that there was engagement with open peer review, utilized open lab notes, plain language, depicted open drafting, was made openly accessible, and used CC-0/BY. The third goal of science/scholarship is that it is efficient; meaning that it connected tools & platforms, there were no publication size restrictions, null result publishing is available, speed of publication is timely, there was the utilization of web standards and IDs, semantic discovery is enabled, there is the ability for re-use, and there are versioning controls. The experimental research workflow mirrors that of traditional research workflow but takes advantage of tools that make the process fully transparent and trackable by an outsider. It is clear from the research that these two have conducted that the workflows for research are being modified and changed depending upon the tools employed. This is a significant development for collection and scholarly communication librarians to be aware of as we consider what aspects of the research lifecycle we want to capture to make readily available from our platforms. Bosman and Kramer's research indicates that we have a whole new world of scholarship practice to consider capturing and retaining in order to provide a transparent and full-fledged depiction of the scholarly record in the twenty-first century.

OpenAPC

<http://treemaps.intact-project.org/apcdata/openapc/#publisher/period=2011>

In Germany, the major research institutions have come together to develop INTACT, a website to provide the transparent infrastructure for article charges. INTACT is made up of three initiatives, Open APC, ESAC, and OA Analytics. The analytics initiative is still in development whereas ESAC (Efficiency and Standards for Article Charges) is well developed outlining best practices for Open Access management around APCs. ESAC, which was begun in 2014, has grown substantially in the past two years and has most recently announced a "Joint Understanding of Offsetting," which depicts pilot offset deals or "Open Access big deals." The intent of these deals is to help with the transition from subscription models into fully Open Access business models. The third initiative, Open APC depicts in graphical form the APC charges paid by over thirty research institutions in Germany. The visualization shows to which publishers and journals the APCs were paid from 2005 to present as well as cost figures given in Euros. The APC Treemap Visualisations are rather fascinating to delve into to find trends and developments. It is impressive to see how great APC adoption has been in Germany over the past three years. For instance, in 2005, the Max Planck Institute was the primary user of APCs and they paid for content from Springer titles. In 2006, IOP (Institute of Physics) is added to the mix of publishers/publications supported by IOP but still just from Max Planck. PLoS (Public Library of Science) does not show up until 2008 but quickly grows by over 18%

in 2009 to become the leading publisher paid APCs by 2011-2012 and stays in the top three up through the current year. Elsevier does not become a major player in the APC market in Germany until 2013 but then exceeds both Springer and PLoS who had been around much longer. Max Planck is the main APC user until 2009 when Munich LMU joins them. By 2012, there are ten institutions in Germany providing APC support for their institutions' scholarship. By 2015 there are 29 institutions and a single centralized fund for APC support. The growth over the past decade and prior to the forthcoming mandate from the European Union (set to be mandatory on all publicly funded research by 2020) is rather impressive. The most fantastic thing about Open APC is that the datasets and structure is fully available on GitHub for other libraries and consortia to replicate similar studies once the data is gathered. When considering what this type of study may look like for varying consortia in the United States and other parts of world, it will help to develop the better economic value and models to be attributed to publication of scholarship worldwide.

SciHub

SciHub: Domain continues to change. Atlantic news item here: <http://www.theatlantic.com/technology/archive/2016/02/the-research-pirates-of-the-dark-web/461829/>

Bastian Greshake's research here: <https://thewinnower.com/papers/4715-correlating-the-sci-hub-data-with-world-bank-indicators-and-identifying-academic-use>

There has been a significant amount written about SciHub, the legality of the site, and concerns over the pirating of commercial content on various academic blogs and through numerous web news channels. What can definitely be said about SciHub is that it is creating a very public focus and debate over what the costs of scholarship should be. In addition to standard scholarly publications such as *Science Magazine* and *Nature*, this story has been covered by *The New York Times, The Washington Post*, and *The Atlantic*. After 20 years of academic consternation and debate, the profit margins of scholarly publishers are being given full consideration beyond academia. This is a game-changer for traditional scholarly publishing access and use. There are already calls being made on some blogs for the use of direct reader technology that would work more directly with article reader sites like ReadCube for the provision and access to content. Like Napster disrupted and changed the model of access to music delivery and purchase, SciHub is likely to bring about a similar change to academic published content. As noted above, part of the change will occur as more transparency is brought to the research and publication cycle overall though tools and mechanisms that make science/scholarship more available. Publishers are also seeing significant revenue increases through the use of licensing single article and single entities through document delivery sites and through the payment of Copyright Clearance Center (CCC) fees. At this point, publishers appear to be quite open to new opportunities to expand dissemination through reader sites given the number of publishers providing content to ReadCube and through CCC document delivery. One indication of how likely these models may take hold are the investigations undertaken by Bastian Greshake, a German PhD candidate. Greshake has performed two analyses of the SciHub data depicting countries of use of the content and both of these studies are quite interesting. During the initial study, he focused on World Bank indicators and how these correlate to download rate. From this correlation, he noted downloads occurred mostly Monday-Friday between the hours of 9 AM and 5 PM normalized for different time zones. He discovered through this correlation that the higher Gross Domestic Product (GDP) means greater downloads as does the higher number of people online tended to mean more downloads. In a follow-up study posted May 23, 2016, Greshake noted approximately 5,800 University/College IP ranges from around the globe had been used to download articles from SciHub. Per his calculations, academic use was between 8% to 10% during a standard work week with a drop off on the weekends and holidays.

Given Greshake's research combined with the continued use of SciHub, it is not beyond reason to see scholarly publishers rolling out more article level purchasing options in the future. There continues to be a growing concern and call for more open science/scholarship. The tracking performed by Germany indicates some rather substantial growth in Open Access publication within that country. The tipping point appears to be looming large on our horizon. The funding models and practices of obtaining and collecting the scholarly record are not likely to change drastically overnight or even from one year to the next but rather continue to be incremental ripples radiating out from the traditional business models developed throughout the last century. ■