ARTICLE EXCERPTED FROM:

INFORMATION STANDARDS QUARTERLY
SUMMER 2014 | VOL 26 | ISSUE 2 | ISSN 1041-0031

ISQ

TOPIC

OPEN ACCESS INFRASTRUCTURE

NISO
How the information world CONNECTS

David Ross

A Publisher's Perspective on the Challenges of Open Access

DAVID ROSS

Open access (OA) publishing and archiving of academic research is becoming an important part of the scholarly communication process. This paper provides a publisher's perspective on the challenges faced developing effective infrastructure in response to this evolving, competitive landscape. More specifically, the paper offers the perspective of what has commonly become known as a "traditional subscription publisher." The views offered are based on first hand experiences of one of these publishers, SAGE, and while many of the issues identified will be common to other traditional subscription publishers it does not purport to be representative of the entire industry. It is an industry that has many players with very different levels of engagement with open access.

Publishers are not a homogenous mass. Scopus indexes over 26,000 academic, peer-reviewed journals from more than 5,000 international publishers. Although the market is dominated by large publishers with portfolios running into the hundreds, if not thousands, of titles, there is a very long tail of smaller operations. These publishers vary significantly in philosophy and corporate structure: from commercial to not-for-profit; university presses to multi-nationals; independents to corporate behemoths; august institutions with hundreds of years of history to relatively new entrants. For those opting to engage with open access, the configuration of their systems and their ability to manage open access publication will vary widely. This paper attempts to present an overview of some of the key challenges in developing open access infrastructures that are common to many but are certainly not universal.

Uncertain Legislative Framework

Academic publishers work in a global environment and their author base is international. Whilst the well-established western markets of Europe and North America still dominate, the emerging economies—led by China and India—are contributing an ever increasing proportion of the research output. As a result, the plethora of national funder mandates provides a very challenging environment for publishers to work within. In addition to these government orders, the numerous private funding agencies have their own OA requirements.

As of June 2014, ROARMAP (Registry of Open Access Mandatory Archiving Policies), as shown in Table 1, lists 466 mandates with a further 27 proposed.

	EXISTING	PROPOSED
INSTITUTIONAL MANDATES	218	7
SUB-INSTITUTIONAL MANDATES	44	4
MULTI-INSTITUTIONAL MANDATES	9	5
FUNDER MANDATES	90	10
THESIS MANDATES	114	

Table 1: ROARMAP OA Mandates *(Source: http://roarmap.eprints.org/)*

Historically these have generally required green OA archiving of research but more recently mandates that make provision for gold OA solutions have also begun to be rolled out. The highest profile has been the Research Councils UK (RCUK) mandate that came into effect on April 1, 2013, but the German Research Foundation (DFG) policy provides funds for pure Gold OA publishing and allows direct grants to be used for Hybrid Gold, and the Dutch funding council has also suggested that they are considering some form of gold mandate.

In the US, there are no less than three initiatives on the table: the outcome of the Office of Science and Technology Policy (OSTP) directive, the Fair Access to Science and Technology Research (FASTR) Act, and the Frontiers, Innovation, Research Science, and Technology (FIRST) Act.

Just recently, on May 15, the National Natural Science Foundation of China, one of the country's major basic-science funding agencies, and the Chinese Academy of Sciences, which funds research at more than 100 institutions, entered the fray with the first major green deposit mandate for China.

Open access has shifted from being a bottom-up, scholar-led movement to top-down, funder-led. But it is not the intention here to examine the relative strengths and weakness of any approach. It is to make the point that the uncertain and ever-shifting global framework presents publishers with unique challenges with respect to long-term strategic planning, shorter-term policy decisions, and the development of infrastructure and workflow solutions to support these.

SAGE, like most publishers, strives to enable author compliance and welcomes well thought-out national mandates with reasonable embargo periods on the availability of the version of record. However, while there is a considerable amount of overlap between these mandates, there are also significant variations in conditions: most often their deposit criteria, embargo periods, and preferred license. Keeping abreast of the evolving framework poses its own challenge and publisher policy changes are often required to reflect these. As an example, in 2013 SAGE adopted one of the most liberal policies with regard to the authors accepted manuscript (AAM), allowing authors to post this in an institutional repository or their personal website immediately, with no embargo. This makes articles published by SAGE compliant with all mandates that have requirements for the AAM but to enable this we were required to consult with all our publishing partners before doing so. (SAGE Publishes on behalf of almost 300 learned societies, associations, and institutes.) It then necessitated alterations to author publishing agreements for all of our 700 plus journals.

Collectively, the industry has shown itself to be willing to engage and seek solutions to these challenges. It instigated the CHORUS project in the US as a possible solution to address the specific request by the OSTP for federal funding agencies to put forward open access solutions to make research derived from their funding public. The project was set up based on CrossRef's FundRef service and CrossRef itself was an industry-funded organization formed specifically to address the need to develop industry-wide standards and provide some infrastructure, originally in relation to DOI technology linking scholarly references. As the global mandate picture develops, it is likely that more initiatives such as this will be required.

The challenge is second guessing future requirements when developing systems and policies and helping authors navigate their way through what can be a complicated and confusing landscape.

Education and Compliance

At a recent workshop, a librarian outlined the problem of an author who had to satisfy five different mandates: Higher Education Funding Council for England (HEFCE), Research Council UK (RCUK), a private foundation, the publisher, and his institution. For an author, interpreting all these is difficult enough; actually ensuring compliance is even more so as there is no silver bullet that will

CONTINUED »

satisfy all five mandates, even if the most liberal embargo period was in place. Authors need assistance and publishers, as well as librarians, have a role to play in both explaining the situation and enabling them to comply. It is likely that the authors or the institutions themselves (it is after all the institution that signs the grant agreement in the majority of cases) will ultimately be responsible for compliance and suffer any consequence of non-compliance, but publishers are being asked to intervene. Initiatives such as SHERPA/FACT, an author/funder compliance tool for RCUK and Wellcome Trust, which relies on publishers' data, may go part of the way in providing a solution at the national level—but there is no global system under development to mirror it worldwide. Calls have been made to make mandates machine readable, to enable automated compliance verification, and in particular to provide accurate solutions for multi-funder cases, but these have yet to be answered.

Another example is that, to ease compliance, the recent mandate announced by the four UK higher education funding bodies will quite possibly result in publishers having to develop entirely new workflows to enable the automatic deposit of AAMs, much in the same way they developed systems to automatically deposit National Institute of Health (NIH) funded papers in PubMed Central (PMC). The mandate requires that for outputs to be eligible for submission to the next UK Research Excellence Framework (REF), the AAMs be deposited in a university institutional repository (IR) at the point of acceptance. UK higher education institutions generally have no comprehensive tool to identify and collect all their published research and metadata, let alone the ability to flag articles at the point of acceptance. While it will be incumbent on authors to work directly with their institutions, it is likely publishers will have a role to play in assisting them to satisfy this REF OA requirement. At present, different stakeholders are developing individual solutions. Coordination and cooperation are required and standardized solutions need to be developed.

Systems and Process

For hundreds of years publishers have operated journal-level workflows. Although the advent of online publication began a shift to article-level workflows, open access publication has accelerated this change. Truly continuous, open access publications operate more or less solely at an article level (with some exceptions in title-level indexing requirements).

There is a need for publishers to interact with authors as paying customers in a way they have not done before.

To facilitate this, publishers have developed new processes to accommodate the needs of open access publishing, often using manual or semi-automated work-arounds in the early stages. Enabling article deposit of NIH-funded papers in PMC, altering production and hosting processes, building systems for authors to pick their preferred Creative Commons license, and developing article processing charge (APC) collection interfaces are just a handful of examples.

One of the challenges for established publishers has been that, generally, they operate using legacy systems designed to service journal publications under the subscription model and these systems are generally not fit for OA purposes. New systems are required in addition to existing ones—the subscription business has not gone away and is not going to in the foreseeable future. Like all organizations, publishers have a multitude of strategic objectives and have to prioritize where they invest their resources. Couple that with the fact that, outside of biomedicine, revenues generated by open access are quite modest and you find the development of systems to facilitate a more streamlined approach to OA are often not deemed business critical.

Article Processing Charge (APC) Collection

A key pain point has been the administration of APCs. Journal publishers are configured to transact large payments with libraries in annual cycles for subscriptions. While those that operate parallel book programs may have some direct interactions with individuals as customers, third-party booksellers handle the majority of financial transactions even in

that sector. There is now a need with OA to quickly process thousands of payments from individual authors and their institutions, and little or no experience by publishers in handling such transactions.

Automated systems need to be developed to take payment by credit card, issue invoices where needed, and apply VAT to European customers only (an issue exacerbated by frequently changing European Union VAT rules). Multiple prices, discounts, split payments, waivers, and currencies have to be handled; transactions and information logged; and reports produced. Institutions which have OA deals in place—some of which are prepaid, some of which are negotiated discount deals—operate under different rules. The institutional approval processes for the use of OA funds are often idiosyncratic but still need be adhered to, and detailed receipts must be issued and regular reports provided. Internally, APCs must be allocated correctly and credit control rules adapted to reflect the much smaller invoice amounts being dealt with.

None of these issues are insurmountable and numerous third-party providers have stepped into the space—from new entrants such as Open Access Key (OAK) to existing intermediaries such as EBSCO, SWETS, and the Copyright Clearance Center.(CCC)—but APC handling requirements have required publishers to invest considerable funds and time in developing modified processes, in addition to existing operations, even when partnering with a relevant intermediary.

Finally, there are often unforeseen effects. For example, with direct payment by authors come greater customer expectations. This direct B2C (business to customer) transaction is seen by some as one of the key drivers of a functional APC market, but it also has implications for customer service functions, expected speed of publication, and additional author services. There is a need for publishers to interact with authors as paying customers in a way they have not done before.

Once again, these new functions are additional to existing operations. Although our interactions with authors may be evolving, SAGE believes the library will continue to be the main transactional partner in the future. The way that libraries are taking on the administration of open access demonstrates this is likely to be the case even in an OA environment.

Licensing

Under the open access publishing model authors are often allowed a choice of a Creative Commons (CC) license, or something broadly similar. Until now publishers have generally operated a single common license for all the research they publish, whether it be through an assignment of copyright or the granting of an exclusive license to publish. There have always been some occasional exceptions, such as those covering US government employees, but the open access publishing model, with a choice of license type, fundamentally changes the legal relationship between the publisher and author.

Again, the intention here is not to debate the merits of CC licenses but to point out that infrastructure implications are not insignificant. Until now, the publisher has been the administrator—and defender—of the copyright in academic research works. Under a CC BY license, that responsibility remains with the authors themselves whereas some derivates, such as CC BY-NC, require the publishers to retain some limited capacity in that role. Systems have had to be developed to automatically recognize license types, add them to article metadata, display them correctly with the associated article, apply the correct permissions criteria, and record the terms in a contracts database for future reference.

Although this task has been made easier by the existence of standard contract templates as established by Creative Commons, the true long-term implications of a large-scale shift of copyright administration to the author has yet to be felt.

Standards and Identifiers

Possibly the biggest hurdle to developing scalable and interoperable systems in any industry is the development and implementation of common standards. It is no different in open access publishing. A great deal of progress has been made in some quarters but little in others.

As an illustration, one of the key problems faced by many in the industry is the inability to automatically identify an author's institution. As mandates proliferate and more and more institutional OA funds are set up with different business rules, the requirement to identify an author's institution to enable automation is becoming key. Returning to the UK REF OA mandate as a case in point, this requires that the AAM of almost

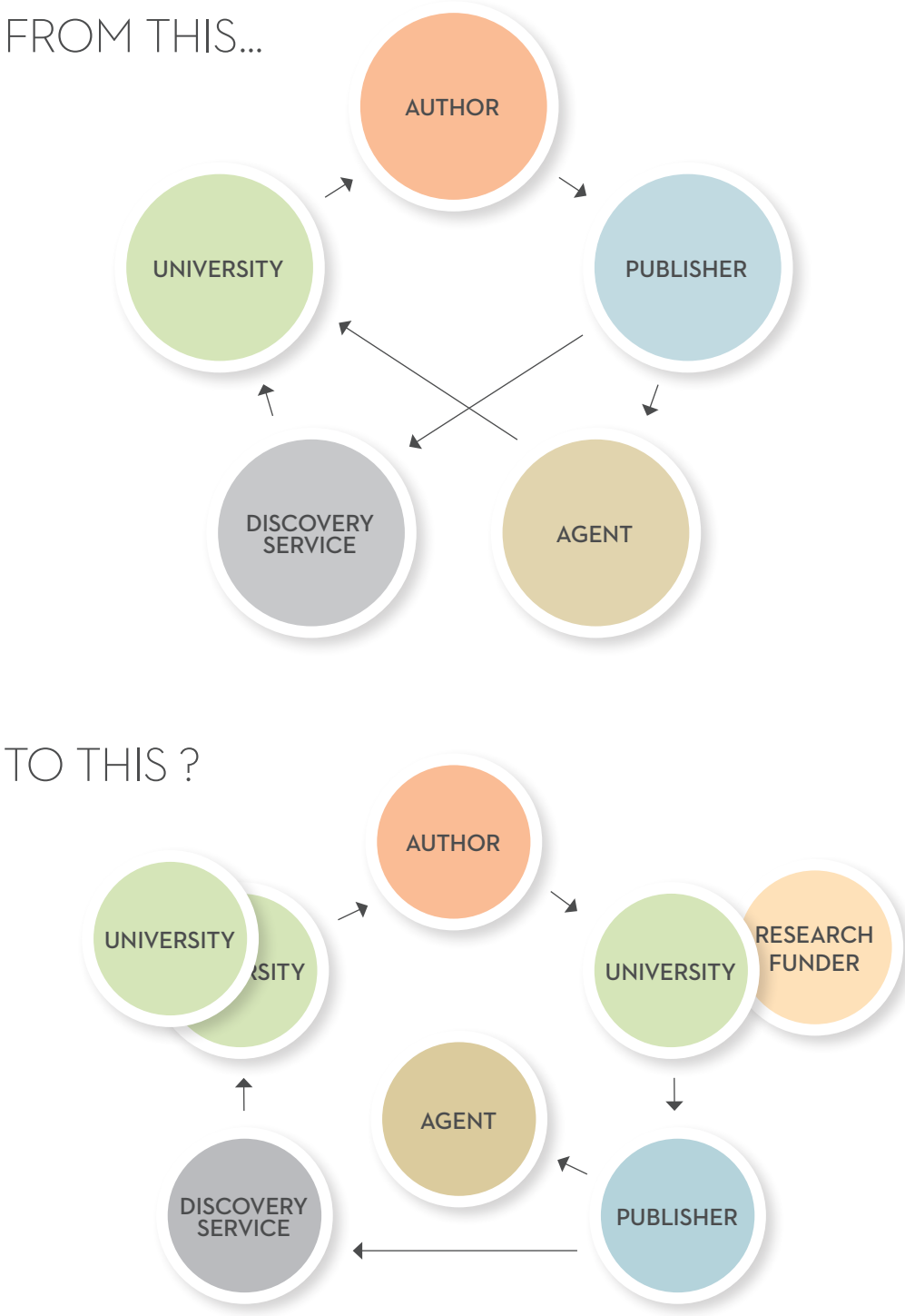

Figure 1. The Changing Roles of Stakeholders in the Information Chain
(Reproduced courtesy of Neil Jacobs, Head of Scholarly Communications Support, Jisc)

every paper written in a UK university is deposited in the relevant institutional repository. The first step toward that is obviously that an institution is alerted to when a paper has been accepted and this can only be done automatically if there are unique, globally recognized standards for identifying higher education institutions—perhaps even down to departmental level. While commercial initiatives such as Ringgold and nascent collaborations such as the Consortia Advancing Standards in Research Administration Information (CASRAI) may provide part of the solution, coordination is required by all stakeholders. The new ISO International Standard Name Identifier (ISNI) standard (ISO 27730) is promising, but is still in the early stages of recognition and adoption. Though author identifiers such as ORCID may one day be part of the solution, it is generally accepted it will be many years until a critical mass of the research community is indexed and registration becomes the norm.

This is only one example of the myriad of new interactions that will be required to develop global, industry-wide scalable and interoperable systems. There will be numerous new interfaces between publishers, authors, institutions, funders, intermediaries, and third-party vendors. Figure 1 illustrates how the relationships between the various stakeholders in the scholarly communication system are changing. It describes movement from a relatively stable environment with well-defined roles for all in the chain—publishers, institutions, authors, agents, discovery agents—to something that is much less clear. There is a new role for funders and several new roles for institutions, not just through paying for open access but also in monitoring and compliance. All these new interactions will require universal standards and identifiers in order for workable APIs to be developed.

There are many initiatives under way that form part of the picture. In addition to those mentioned above, NISO itself has created the Open Access Metadata and Indicators Working Group (renamed the Access and License Indicators Working Group), which is currently finalizing its recommendations following public consultations (see article on page 35). Jisc has started exploring the development of a managed shared service, Jisc Monitor, which might support UK institutions, central to which will be the adoption of standards to enable the interoperability required.

Legacy systems currently in use by all the stakeholders complicate matters further. Publishers' internal systems often have difficulty interfacing with one another and the complexity of enabling these to interface with an entirely new set of external systems is not to be underestimated. Consider also that publishers generally depend on a variety of third-party vendors for key parts of their workflow. Manuscript processing systems and hosting platforms are generally contracted out and these have their own limitations including being generally built for journal-level workflows. These vendors have multiple customers and multiple, sometimes conflicting, demands, which means that they are not always the most flexible or swift at adapting.

One Size Does Not Fit All

The open access publishing market as driven by APCs has predominately developed in the biomedical market, but it has always been accepted that humanities and social sciences (HSS) would present their own challenges. A recent study commissioned by the British Academy, *Open Access Journals in the Humanities and Social Science*, went further and suggested that the market should not so much be viewed as STM vs. HSS, but rather biomedicine vs. the rest. Certainly OA in the humanities, where the monograph is the main convey or of information, faces its own challenges, but many other social science disciplines, where the research article is still the main vehicle, face their own particular problems, mainly due to the relative paucity of funding. In many HSS disciplines, the journal itself serves a very different purpose than in STM. Selection mechanisms are different and the necessity for expert opinion to confer authority on scholarly work that deals with concepts and ideas rather than empirical data requires different approaches.

This last point is made to illustrate that the development of the open access market is not evenly distributed, by geography or subject. For large global publishers that cover a range of disciplines this presents yet another level of complexity as they attempt to find viable long-term solutions that satisfy all the requirements of all major stakeholders: authors, editors, societies, universities, and funders.

| IP | doi: 10.3789/isqv26no2.2014.04

DAVID ROSS (david.ross@sagepub.co.uk) is Executive Publisher of Open Access for SAGE Publishing (www.sagepub.com) where he oversee a team that manages SAGE's flagship Open Access journals and develops systems, process, and policies to support their OA publishing.

CONTINUED »

CAS [Chinese Academy of Science] Issues Open Access Policy
http://english.cas.cn/Ne/CASE/201405/t20140516_121037.shtml

CHORUS (Clearinghouse for the Open Research of the United States)
http://chorusaccess.org/

Consortia Advancing Standards in Research Administration Information (CASRAI)
http://casrai.org/

Copyright Clearance Center Launches RightsLink for Open Access
http://goodereader.com/blog/e-book-news/copyright-clearance-center-launches-rightslink-for-open-access

Creative Commons
http://creativecommons.org/

CrossRef's FundRef
http://www.crossref.org/fundref/

Darley, Rebecca, Daniel Reynolds, and Chris Wickham. *Open Access Journals in the Humanities and Social Science.* The British Academy, 2014.
http://www.britac.ac.uk/templates/asset-relay.cfm?frmAssetFileID=13584

Fair Access to Science and Technology Research (FASTR) Act
https://beta.congress.gov/bill/113th-congress/house-bill/708

Frontiers, Innovation, Research Science, and Technology (FIRST) Act
https://beta.congress.gov/bill/113th-congress/house-bill/4186

Higher Education Funding Council for England (HEFCE)
http://www.hefce.ac.uk

International Standard Name Identifier (ISNI)
http://www.isni.org

Jisc Monitor
https://www.jisc-collections.ac.uk/Jisc-Monitor/

New Policy for open access in the post-2014 Research Excellence Framework: HEFCE Press Release, March 2014
http://www.hefce.ac.uk/news/newsarchive/2014/news86805.html

NISO Access and License Indicators Working Group
http://www.niso.org/workrooms/ali/

Office of Science and Technology Policy (OSTP) Memorandum: Increasing Access to the Results of Federally Funded Scientific Research
http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Open Access and research funding by the DFG (German Research Foundation)
http://www.dfg.de/en/magazine/spotlight/open_access/

Open Access Key (OAK)
https://www.openaccesskey.com

***Open where possible, protected where needed: NWO and Open Access.* Netherlands Organisation for Scientific Research, November 2012.**
http://www.nwo.nl/binaries/content/assets/nwo/documents/nwo/open-access-flyer-2012-def-eng-scherm.pdf

ORCID
http://orcid.org/

PubMed Central (PMC)
http://www.ncbi.nlm.nih.gov/pmc/

***The Research & Innovation Performance of the G20.* Thomson Reuters, March 2014.**
http://sciencewatch.com/sites/sw/files/images/basic/research-innovation-g20.pdf

RCUK (Research Council UK) Policy on Open Access
http://www.rcuk.ac.uk/research/openaccess/policy/

Ringgold
http://www.ringgold.com

ROARMAP (Registry of Open Access Mandatory Archiving Policies)
http://roarmap.eprints.org/

SAGE research funder mandate compliance information for authors
http://www.sagepub.com/author-info.sp

Scopus
http://www.elsevier.com/online-tools/scopus

SHERPA/FACT (Funders & Authors Compliance Tool)
http://www.sherpa.ac.uk/fact/

Wellcome Trust Open Access Policy
http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Open-access/

RELEVANT LINKS