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


Martin Moyle


Catherine Sharp


Alan Bracey

The Role of Standards in the Management of Open Access Research Publications:

A Research Library Perspective

MARTIN MOYLE, CATHERINE SHARP, AND ALAN BRACEY

UCL (University College London) is a multidisciplinary, research-led institution with approximately 5,000 staff and postdoctoral researchers and 4,500 research students. As London's Global University, UCL engages with the spectrum of research subjects, from arts and humanities to basic and applied sciences and medicine. UCL embraces open access (OA), supporting both the Gold and Green routes to OA. Academic freedom is a cornerstone of UCL's OA Publications Policy: UCL researchers are free to determine where to publish, how much to publish, and how often to publish.

OA support at UCL

Open access at UCL is championed by the Vice-Provost (Research) and strategy is led by a Publications Board of senior academics and administrators. Operational responsibility for OA resides with UCL Library Services. The Library OA Team is responsible for UCL Discovery, UCL's Institutional Repository, ensuring that Green OA is supported efficiently and legally. The Team also manages Gold OA, with responsibility for overseeing UCL's Gold OA publications budget and for helping researchers to align their publication practices with the OA policies of funders, in particular those of the Wellcome Trust, Research Councils UK (RCUK), and, recently, the UK Higher Education Funding Council for England (HEFCE).

CONTINUED »

Key OA activities

UCL Library Services supports OA at UCL through three key activities:

- Arranging the payment of Article Processing Charges (APCs), including the negotiation and maintenance of agreements with publishers that secure value for money and/or administrative efficiencies for UCL.
- Storing and providing access to Green OA. UCL encourages self-upload of research by authors; on receipt, permissions are checked and embargoes applied before access is enabled.
- Collecting, analyzing, recording, and reporting management data.

OA research publication: stakeholders and standards

Standards are particularly important in multi-stakeholder environments in which different parties need to share information with consistency. Open Access research publication is one such environment. The primary stakeholders in this context include authors, publishers, funders, and a range of administrative units within higher education institutions (HEIs) supporting OA management and research administration. The main stakeholders are supported by several service providers and agencies: for instance, vendors of research information systems and manuscript submission systems, registration agencies, open source developers and relevant organizations such as SHERPA and Creative Commons.

To illustrate some of the opportunities for interplay between stakeholders and standards in the area of OA research publication management, four brief case studies are outlined. The first considers the position of the author at publication; the following studies look more closely at the three "key OA activities" identified above.

1 *Author obligations*

At publication, authors face the task of identifying and complying with a range of requirements imposed by their publishers, funders, and institutions.

At the earliest stages of publication, authors need to work with their publisher's submission system, often proprietary and custom-made. They must provide required author identifiers, which may include ORCIDs or proprietary identifiers, if known. They must give their affiliation—not without difficulty at UCL, whose many postgraduate institutes and interdisciplinary research centers offer a variety of possibilities for diversity of attribution. Publishers may ask authors for a code or institutional affiliation to establish that an institutional prepayment scheme for Gold OA applies; authors need to be aware that these publisher-institution relationships exist, and to know how to find out whether their institution has an agreement with their publisher. A title-specific Copyright Transfer Agreement (CTA) must be endorsed or amended and, ideally, its implications fully understood by the signatory.

To ensure that funder requirements are met, authors need to be aware of all the different funding sources that contributed to the paper, and to understand where to find guidance on those funders' open access policies. This may direct or limit their publication choices: the Wellcome Trust, for example, strongly recommends Gold open access; and many UK institutions have block grants to cover APCs for RCUK-funded papers. Funder requirements also have a bearing on licensing, obliging the authors to assign a particular license (usually CC BY in such cases), sometimes necessitating further negotiation with the publisher. Authors need to know about and act on any stipulations concerning the acknowledgement of funders and grant numbers, either of which may be elicited as free text contributions rather than selected from a controlled list. (Authors can be asked to provide grant details several times over, in fact—in the manuscript, as metadata at final submission, in correspondence with administrators about APC payment, and, later, as part of institutional processes around publication recording and reporting.) Other funder idiosyncrasies, such as RCUK's mandate for a statement about the availability of underlying data, must also be observed at submission.

Researchers must understand and comply with any requirements placed upon them by their institution. Local open access policies may direct their final publication choices and affect their engagement with CTAs and licenses. Typically, the institution requires a record of every publication produced by its researchers; the author must understand how such data is compiled and contribute to its upkeep. Publication harvesting systems are becoming more commonplace, but can be rather "hit and miss" affairs, not least because of author and institutional identification issues mentioned above, as well as disciplinary differences in the coverage of available harvesting sources. An author may have to make a manual intervention to ensure that the local publication record is up to date. Finally, and to close the loop within institutional systems, the correct associations between individual research outputs and local records of the funders and grants under which those outputs were incubated must be recognized and recorded.

2 *Arranging Gold open access*

For HEIs, arranging Gold payment can be a complex process. Funder requirements for Gold OA must be matched to publishers' Gold offerings; authors often need advising on HEI, publisher, and funder policies; and payment arrangements developed for the subscriptions environment do not translate to the payment of article processing charges, necessitating new financial processes.

UCL's Open Access Team arranges APC payments on behalf of authors, who may be addressing the requirements of a funder that strongly encourages Gold open access (the Wellcome Trust, for example). Authors may choose Gold if immediate open dissemination is required. As the rapid growth of Gold offerings from publishers has been an organic process, the level of "openness" offered varies widely—as a result, funder requirements for Gold are not always met by publishers.

When the Team receives a request, they verify an author's affiliation and eligibility to use UCL funding. The verification process is obstructed when up-to-date data is unavailable in local HR or research systems, when an author is affiliated with a number of different institutions, or when a non-UCL author requests Gold open access on behalf of a UCL co-author. Widespread adoption of ORCID identifiers would undoubtedly contribute to a more efficient process for confirming eligibility. Bibliographic and funding details are obtained and recorded in the OA Funding Team's database, and the Team explicates funder policies to authors and clarifies if, and how, a journal's Gold option will comply. Many publishers offer Wellcome Trust- and RCUK-funded authors a choice of license, even though only CC BY is acceptable to the funder. The OA Funding Team directs authors to the type of license required by their funder, and troubleshoots cases where a non-compliant license has been chosen. Differing funder and publisher policies cause considerable confusion to authors and institutions. Funding details are an essential prerequisite for accurate assistance, but details provided to the Team by authors do not always match those given to the publisher, which may differ again from those acknowledged on the paper. The widespread adoption by publishers of a system for collecting and standardizing author's funding data (FundRef, for example) would greatly improve Gold open access workflows for all stakeholders.

Despite the existence of some relevant standards, supporting OA—certainly at UCL—depends uncomfortably on manual intervention and pragmatics.

At UCL, Gold payments are made either through publisher prepayment schemes or by invoice. Publisher schemes are preferred for the efficiency savings gained from not paying hundreds of individual invoices. The OA Funding Team records all payment, bibliographic, compliance, and deposit data centrally, updating records as transactions progress. Manual checks are performed to confirm whether funds have been released by University Finance, whether the publication has been made open access, whether the correct license type has been applied, whether funding is acknowledged (RCUK), and if the paper includes a statement on access to underlying research materials (RCUK). The Team also deposits the final PDF in UCL's institutional repository.

The potential for off-the-shelf APC management systems to deliver efficiencies is noted; however, the emerging systems have not yet transcended the complexities of servicing Gold OA on the scale at which the UCL Team operates.

3 *Supporting Green open access*

As with Gold open access, institutions need to be able to give authors accurate advice on how to engage with Green OA. Publisher rights cannot legally be disregarded, and institutions risk reputational damage (and perhaps financial penalties) if such copyrights are consistently breached. Accurate information about applicable Green rights is required at title level at the point of repository deposit. In the absence of standard input from publishers, it is difficult for an aggregator, like SHERPA (whether as SHERPA/RoMEO or SHERPA/FACT), to render accurately all permutations of a publisher's open access policies. The SHERPA APIs are potentially a great adjunct to repository and publication management systems, but, for full confidence in the legality of Green collections, guidance from the SHERPA suite can currently only be regarded as indicative.

CONTINUED »

The UCL Team frequently liaises directly with publishers and authors over the small print of CTAs. Every deposit is checked for legality and embargo periods identified and applied, before it is made openly accessible. Deposits to institutional repositories in the UK will soon increase dramatically in response to HEFCE's Open Access Policy, which requires the Green deposit, at acceptance, of a final manuscript of every article and proceedings paper that is to be eligible for the next Research Evaluation Framework (REF) assessment. To sustain support for Green OA at this level, institutions need reliable systems for the exchange of rights information, and to free them from the need to check every copyright agreement manually, while remaining fully confident that they are correctly observing publishers' rights and embargo rules.

The emphasis of the HEFCE OA Policy on deposit at acceptance also brings some new challenges in metadata management. Records will increasingly enter institutional systems at acceptance, when metadata is skeletal, and will need to be improved after publication so that they are fit to expose later to HEFCE for REF purposes. The publication harvesting systems already in use could help to automate local record enhancement, but their reliability is compromised by the fact that the unambiguous systematic identification of outputs only becomes possible at publication, when DOIs are disclosed (assuming they have been assigned at all). While HEFCE's support for Green OA is welcome, the risks of the duplication of records and an overall degradation of metadata quality in university systems are naturally increased by the terms of its mandate. There is a need to improve the ability of business systems accurately to disambiguate and merge metadata records to avoid such maintenance becoming an extremely time-consuming manual process for authors and/or administrators.

Institutions naturally wish to capitalize on their investment in repositories and one way of doing so is to celebrate the impact of open research. Monitoring the quantity and provenance of full text downloads is an obvious way of tracking impact. Benchmarking such impact across institutions, however, is made more difficult by non-standardized data collection. How much time should elapse between repeat downloads for each to be counted afresh? Are we sure that all crawlers and harvesters are being excluded from the figures? The IRUS-UK service, which aggregates COUNTER-compliant article-level statistics for cross-repository comparison, is a welcome development, and adoption of the same protocols in local repositories would help to boost the accuracy and credibility of repository impact assessment by institutions.

4 *Compliance monitoring and reporting*

Accurate compliance monitoring is a challenge. To begin with, a means of identifying all the outputs associated with the institution and funded by each relevant research funder must be in place. Such a conspectus is difficult to achieve because of the vagaries of author and funder identification in the workflows outlined above, and so efforts to determine compliance are compromised at the outset.

In preparation for compliance reporting, institutions need to assemble bibliographic data about Gold and Green articles, including DOIs and institutional or subject repository identifiers; funder and grant information; information about licenses, acknowledgements, and other required statements; details of deposit in any specified subject repositories; financial information about APCs and their breakdown between funders; and details of acceptance dates, publication dates, and embargo periods. These administrative and bibliographic metadata need to be stored in readiness for report to different funders—who, of course, have different reporting requirements and emphases, not least because their mandates specify different criteria for compliance. Often, the data elements required for reporting are stored across several local systems, particularly where the institution does not have a current research information system (CRIS). Reports for consumption within the institution, meanwhile, also need attention; monitoring the level of researcher engagement with the HEFCE mandate, for example, is set to be of intense importance to heads of research in UK universities in the near future.

Publisher pre-payment systems have some merit in easing the burden of reporting, in that they typically require the publisher to make a periodic disclosure of how the money deposited by the institution has been used. Such reports, however, do not yet conform to any standardized format or content, and the number of schemes on the market is in any case few, so they are only a small, if helpful, part of the overall picture. There is very little escape from manual data collection and assessment in the sphere of compliance. The UCL OA Team spends much time interacting with authors and publishers to collect, verify, and store a wide range of information in support of future reporting needs.

Current standards and services to facilitate OA research publication

The foregoing illustrations highlight the fact that, despite the existence of some relevant standards, supporting OA—certainly at UCL—depends uncomfortably on manual intervention and pragmatics. The illustrations also show that authors shoulder a great deal of responsibility for policy awareness and administrative information provision, areas into which ideally they would not have to be so deeply drawn. It is noteworthy that many of the concerns of the stakeholders in OA research publication, such as the accurate identification of authors, funders, grants, publisher rights, and individual published outputs, are shared. Such repetition of need strengthens the argument for the development and implementation of standards-based interoperability.

Table 1 summarizes the authors' perception of the current availability and utility of standards, and services with the potential for standards-based development and integration, that would facilitate the OA support work outlined above.

CONTINUED »

TABLE 1: RELEVANT OA STANDARDS AND PROCESSES AND CURRENT STATUS

AREA	REQUIREMENT	WHERE ARE WE NOW?
Researcher identification	Unambiguous identification of authors.	ORCID has much promise; it transcends proprietary researcher ID services and tracks individuals independently of institution. To be fully effective, requires adoption by all publishers and all researchers.
Funder identification	Unambiguous identification of the funder(s) of a research output—useful to funders, HEIs, and publishers.	FundRef offers publishers a normalized list for use in submission systems. Would require full take-up by publishers to be effective.
HEI identification	Correct association of authors and papers with parent HEIs, enabling accurate attribution of citations and supporting compliance monitoring.	ISNI standard (ISO 27729) has promise; and it is noted that ISNI and ORCID have recently announced joint agreement. However, ISNI is not currently widely adopted in submission systems, therefore variant nomenclature is in use.
Digital object identification	Unambiguous identification of a published output.	CrossRef is proven and has high levels of buy-in. DOIs only released at publication; would help HEIs immensely if DOIs could be assigned and routinely shared at acceptance.
Bibliographic metadata exchange	Seamless, automatic population of institutional publication systems.	Several metadata standards, not always employed. Each consumer system needs custom solutions per data source.
Administrative metadata exchange	Funder information, acknowledgements, license details, etc.	No standard. Metadata usually provided via ad hoc add-ons to bibliographic metadata supply, or derived manually by administrators.
Journal submission	Handover of final accepted manuscript to publisher.	Systems and workflows vary between publishers.

TABLE 1 CONTINUED »

CONTINUED...

AREA	REQUIREMENT	WHERE ARE WE NOW?
APC processing	Regularized workflows to manage APC transactions, especially financial aspects.	Solutions being developed, but hitherto unsuitable for large HEIs. Current dependency on per-publisher arrangements.
Publisher rights	At-a-glance, ideally machine-readable, and accurate summary of self-archiving and embargo policies, per journal.	SHERPA-RoMEO is a useful dataset; has API; but currently cannot be relied upon to automate self-archiving without risk. Needs full publisher buy-in to a standard set of rights metadata.
Journal compliance with funder OA policies	Help authors and administrators to identify appropriate journal in which to publish.	SHERPA-FACT tool. RCUK and Wellcome Trust only. Not standards-based. Has potential if input/output can be standardized and coverage widened.
Licensing	Funders increasingly specify the assignment of particular licenses. Authors need clarity; publishers need to guide and support author choices.	Creative Commons. Mature concept, increasingly well understood, but not yet fully embedded in publisher workflows.
Repository downloads	HEIs would like to show impact of Green OA.	IRUS UK—aggregation based on COUNTER standard—good model. Similar standards not necessarily applied in local IRs.
Compliance monitoring	Ensure HEIs meet all obligations to research funders.	Need to combine and analyze HEI/author/funder/license data about publications. Some relevant standards in these areas (see above), but adoption too patchy to be dependable.
Compliance reporting	Accountability of HEIs to funders.	Requirements are funder-specific, although with much overlap between funders. Designation of a common framework for core elements of reporting would enable efficiency gains at report-making institutions.

Conclusion

For standards to be useful to any given community, they must be adopted by all its members, and members' business systems must support the exchange of the information that has been standardized. The short case studies given above indicate that the stakeholders within the sphere of OA research publication do not benefit from the effective standardization of even the relatively few key pieces of information that support commonplace OA interactions, let alone the exchange mechanisms that would enable such harmonized information to flow in a timely way between systems. Funders are channeling money into Open Access—for example, RCUK has committed £20 million to OA from the UK research budget in the current financial year—while HEIs are making increasing investment in repository services and OA administrators, who spend their days chasing down information from authors, publishers,

funders, and finance systems, and maintaining and reconciling multiple spreadsheets for myriad different accounting and reporting purposes. Meanwhile, investment is clearly being made into relevant standards and services, but it is uncoordinated. Improvements are incremental and piecemeal, and meaningful, efficiency-delivering integration is frustratingly elusive.

All the stakeholders in OA research publication would benefit from further investment in the definition, refinement, promotion, and integration of relevant standards. There are common needs and shared problems, and there is new money in the ecosystem. It is interesting to reflect on the potential for publisher systems to act as a "hub" for the exchange of a significant quantity of information between the stakeholders in OA research publication. During the publication process, it is conceivable that author IDs, unambiguously-identified HEIs, DOIs, registry-sourced funder and grant details, standardized rights and license information, at-acceptance and at-publication metadata, financial transactional information, and, indeed, the final accepted manuscript itself might seamlessly change hands between publishers, universities, and funders, drawing on registries and related services (and, additionally, leaving authors with more freedom to concentrate on writing). The traditional role of the publisher is challenged by OA, particularly as universities are taking more responsibility for the dissemination and curation of their published assets. New publisher-led initiatives, protocols, and systems to facilitate standards-based communication between the stakeholders in OA research publication might be warranted. Any such developments would certainly be welcomed by staff who are involved in the administration of open access at universities.

| IP | doi: 10.3789/isqv26no2.2014.03

MARTIN MOYLE (m.moyle@ucl.ac.uk; ORCID 0000-0002-2065-0508) is Group Manager, IT and Bibliographic Services, **CATHERINE SHARP** (catherine.sharp@ucl.ac.uk; ORCID 0000-0002-4472-943X) is Open Access Funding Manager, and **ALAN BRACEY** (a.bracey@ucl.ac.uk) is Open Access Compliance Officer—all with UCL Library Services, UCL (University College London), UK.

RELEVANT LINKS

COUNTER Code of Practice for Articles
http://www.projectcounter.org/counterarticles.html

Creative Commons
http://creativecommons.org/

CrossRef
http://www.crossref.org/

FundRef
http://www.crossref.org/fundref/

HEFCE Policy for open access in the post-2014 Research Excellence Framework
http://www.hefce.ac.uk/pubs/year/2014/201407/

IRUS-UK
http://www.irus.mimas.ac.uk/

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

Open Access at UCL
http://www.ucl.ac.uk/library/open-access/

ORCID
http://orcid.org/

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

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

SHERPA-RoMEO
http://www.sherpa.ac.uk/romeo/

UCL Discovery: Unlocking UCL research
http://discovery.ucl.ac.uk/

UCL Publications Policy 2012
http://www.ucl.ac.uk/library/about/strategies-policies/publications-policy

UK Higher Education Funding Council for England (HEFCE) Policy Guide Open Access Research
http://www.hefce.ac.uk/whatwedo/rsrch/rinfrastruct/oa/

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