Managing open access with EPrints software: a case study

Recent additional open access (OA) requirements for publications by authors at UK higher education institutions require amendments to support mechanisms. These additional requirements arose primarily from the Research Councils UK Open Access Policy\(^1\) applicable from April 2013, and the new OA policy for Research Excellence Framework\(^2\) eligibility published in March 2014 and applicable from April 2016.

Further provision also had to be made for compliance with the UK Charities Open Access Fund, the European Union, other funder policies, and internal reporting requirements.

In response, the University of Glasgow has enhanced its OA processes and systems. This case study charts our journey towards managing OA via our EPrints repository. The aim was to consolidate and manage OA information in one central place to increase efficiency of recording, tracking and reporting. We are delighted that considerable time savings and reduction in errors have been achieved by dispensing with spreadsheets to record decisions about OA.

Background

The University of Glasgow has a well-established centralized open access (OA) service. We believe this provides us with economy of scale and delivers the most consistent and efficient service.

The vast majority of the administrative work is done within the library with the aim of minimizing the burden on the author. We estimate this has taken an additional three full-time equivalent staff since 2013. This resource has been taken from existing roles. Even after factoring in efficiency savings from our amended systems and processes, we expect to require additional resource to deal with the projected increased volume of work from 2016 due to changes in UK funder and government reporting requirements.

If decisions on direction are required, we are supported by our Vice-Principal for Research who is an active researcher and user of the OA service. He chairs the University Research Planning and Strategy Committee, a group that can be called on to provide advice or decisions where appropriate. Members of the Committee are also important in ensuring that key messages are disseminated to the wider community. Despite regular meetings, stalls and outreach events, one of our biggest concerns is that not every author will become aware of these OA requirements or the support available. Another big push will take place early 2016 and we expect this to be an ongoing activity.

The OA workflow

The OA workflow at the University of Glasgow is:

- authors – staff or students – are encouraged to e-mail a generic e-mail address (research-openaccess@glasgow.ac.uk) as soon as they are notified of acceptance of their article or conference proceeding. Each enquiry is dealt with on a case-by-case basis. There is a template of standard e-mail replies that can be tailored to most enquiries including:
  - confirming that the library will arrange payment of an invoice
  - asking for the final agreed text if we are not paying for open access but a no additional cost route is available

\(^{1}\) Research Councils UK Open Access Policy
\(^{2}\) Research Excellence Framework
· advice on obtaining a PubMed Identifier
· open access advice for books and book chapters.

· Library staff then enter the details onto the publications repository ‘Enlighten: Publications’

· authors and administrators may also add records direct to the publications repository. All of these records go for review and OA checks by Library staff before they go live

· awards are added from the searchable import from our Research System

· an OA administration category is available to Library staff and is added to each record (Table 1). This categorization is open for change, e.g. we now think the ‘outwith scope’ category can be dispensed with

· a notes field is used to record the status of the record, e.g. ‘ Asked author for funder award numbers and asked author to confirm that the Biotechnology and Biological Sciences Research Council and their award are acknowledged in the paper.’

The final stage depends on the OA status of the article. For articles identified as:

· green – the final agreed text is requested and the repository team follows this through to complete the record with publication date and release the full text after any embargo

· gold – details are sent to the library Acquisitions and Access team who arrange payment on behalf of the author. The repository team completes the record after publication ensuring the licence type is CC BY where necessary and ensuring the full text is openly available

· outwith scope/no green OA option/no OA option – the repository team complete the metadata for the record.

<table>
<thead>
<tr>
<th>Type open access</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>Article will be made freely available without any additional cost</td>
</tr>
<tr>
<td>Gold</td>
<td>OA costs will be paid from funds administered by the library</td>
</tr>
<tr>
<td>Pending</td>
<td>Decision to be made, e.g. awaiting further information on funding from author or clarification of licence options from publisher</td>
</tr>
<tr>
<td>Outwith scope</td>
<td>The item does not have OA requirements e.g. a book chapter that is not funded by</td>
</tr>
<tr>
<td></td>
<td>Welcome Trust does not have OA requirements</td>
</tr>
<tr>
<td>No green OA option</td>
<td>There is no OA option without additional cost</td>
</tr>
<tr>
<td>No OA option</td>
<td>No OA options at all available for this journal</td>
</tr>
<tr>
<td>Other</td>
<td>For use where other categories do not explain. Rarely used and may be removed</td>
</tr>
</tbody>
</table>

Table 1. Open access administrative categories at The University of Glasgow Library

Reports are generated direct from the publications repository and can be used for monitoring increase in OA engagement, identifying opportunities for improving the support provided, reporting to funders, and contributing to the Jisc Total Cost of Ownership project (Figure 1).

Specification

In 2013 the University had been researching and collating new OA metadata requirements to deliver a comprehensive system that would provide users with all the information they needed, avoid duplication of effort, and minimize errors. We wanted all the information in one place otherwise users would create other places to keep what they needed.
We created a metadata specification after reviewing RCUK (Research Councils UK), REF (Research Excellence Framework), EU and internal requirements. During 2013-4 we publicized this to staff at HEIs (higher education institutions) who administer open access via mail lists such as ARMA (the Association of Research Managers and Administrators) and The United Kingdom Council of Research Repositories, and meetings and workshops. We consulted users of other systems such as PURE and Converis, users who did not have any systems, and other EPrints sites. We wanted to avoid duplication of effort for HEIs in determining what was required and finding systems solutions, and believed our draft specification would be more robust if it were comprehensive, transferable and could be standardized. Our specification was updated as new information became available, notably the RIOXX initiative, whereby a beta version of the specification for RCUK compliance was released in June 2014. Once we reviewed this specification, we understood that it mainly consisted of fields we already held on our system.

Around September 2014 we decided to place amendments to our generic community specification on hold as a new initiative had arisen to look at national standards for OA metadata. The CASRAI (Consortia Advancing Standards in Research Administration Information) UK Open Access Working Group would attempt to co-ordinate this standard. This provided a better structure and formal stakeholder engagement including funders and HEIs. The group took the metadata specification we had provided along with RCUK, REF, EU and many other requirements as inputs to draft standard definitions for open access.

The RIOXX initiative was waiting for new NISO licence definitions to be released and once this information was available, a first formal version of the RIOXX specification was released in January 2015.

In July 2015 the first tranche of CASRAI UK definitions focusing on APCs (article processing charges) was released for comment.

We started a Jisc Pathfinder project ‘End-to-End Open Access Process Review and Improvements’ in June 2014 and this helped facilitate the discussion around metadata requirements.
System development

Prior to the RCUK Open Access Policy, we already had many fields in our system to manage OA. These included full-text status to indicate if an article was fully OA, restricted (often due to become fully OA but currently restricted due to publisher embargo) or not available in our repository. When the RCUK Policy was issued, we added additional fields including an administrative field to record decisions about OA. These decisions included recording whether the article was green, gold, had no OA option at all, or pending – for example if we were awaiting funder information from the author, or had asked a publisher to clarify their policy.

Timing was a concern as lead time for system development, testing and embedding was diminishing and, like other HEIs, we were becoming anxious to have systems in place to capture the information that would need to be reported. There were a number of pieces of the jigsaw missing; we did not know what they would look like or whether they would fit. For example, we understood that a plug-in for EPrints for RCUK compliance would be available but were unsure as to when that would be and what the tool would deliver – would it simply add new fields or would it produce a formatted report to facilitate our monitoring and return?

In August 2014 conversations were ongoing with Jisc Monitor project representatives who started to look at a simple APC specification for reporting to funders and benchmarking OA costs.

In autumn 2014 our return to RCUK was assembled with considerable manual manipulation of data from several places including our EPrints system, Finance System, and spreadsheets. We knew this was inefficient.

The EPrints plug-in for RCUK compliance, developed from the RIOXX initiative, was provided for testing in January 2015. As we linked this tab to fields which were mostly pre-existing in our system, we realized it was not sufficiently comprehensive for our needs. For example, it did not include a field to indicate if the author had acknowledged research materials in the paper – a requirement stated in section 3.3 of the RCUK Open Access Policy. The report provided was solely for checking compliance with the profile, whilst we needed other reports to assist us with identifying gaps and errors in our data. However, given that RCUK endorsed the tool as an indicator of compliance, we deployed it live.

We then set about adding additional fields, mostly financial, to our EPrints system. Most of the information was copied manually from our Finance System as there was no automatic interface between the two. Creation of an automated linking between the two systems was not pursued as there was no unique identifier reliably available in both systems. We provided a test login to our partner site, Southampton, and received useful feedback.

Our OA and RCUK functionality was made live on 27 March 2015, providing a mechanism to gather everything we needed for OA within our EPrints system, as well as some basic reporting (see Figure 2) that we can use to monitor our compliance and provide information to colleges and funders.

We are delighted that this means we have been able to dispense with spreadsheets for data collection. This reduces duplication of data entry, potential for error, and annoyance and delay at being locked out of the spreadsheet whilst others edit it.

It became clearer early in 2015 that Jisc would be providing tools for EPrints to facilitate future REF compliance, so we then knew we did not need to develop full functionality to cover REF OA compliance but could wait for that to be provided for testing.
We subsequently amended our terminology and fields slightly to align with the Jisc format that was recommended for reporting to the UK Charities Open Access Fund (COAF) in October 2014. A variation on this format has subsequently been recommended for reporting to RCUK.

On 9 June 2015, after some live usage at Glasgow, we were comfortable enough to share the code for our EPrints OA fields and reporting on GitHub. Anyone interested in testing in their own system could find an EPrints plug-in file to install, or create their own version, from the code.

We sent test extracts of data downloaded from EPrints to Jisc representatives who confirmed that our report emulated the format they had been working on for the Total Cost of Ownership Project. See Figure 3a and 3b for an example extract from our EPrints repository to the required format.

As noted above, the same format satisfies the formal reporting requirements of RCUK and the COAF, albeit with some minor amends as each currently has a slightly different format. In response, Jisc updated the advice they give to note that downloads from systems are possible. https://www.jisc-collections.ac.uk/Jisc-Monitor/APC-data-collection/Jisc-APC-template-FAQ/

In November 2015 a test version of the REF OA plug-in for EPrints was received from Jisc. The plug-in allows organizations to categorize and evidence exceptions to the policy. We deployed this in a test environment in November and installed the released version in our live repository in January 2016. See Figure 4 for screen shots for this new functionality. We are monitoring the information captured via this tool and will provide feedback to Jisc on the practicalities of its use.

Next steps

We continue to encourage co-operative working and give live demos of our work, so far at events such as the EPrints UK User Group and on an ad-hoc basis with individual sites.

Activities include:

- working with a site in Germany that is looking at our code as they start to do similar work. We have also received some visualisation code from them and we hope to work together to improve these offerings to the EPrints community in the near future
- integrating the REF OA tool with the standard REF functionality for future REF exercises. This functionality has still to be fully specified and delivered. The exercise is not due until 2020 or beyond. We must, however, capture evidence of exceptions to the OA policy from 1 April 2016 in accordance with the REF Open Access Policy
- we are testing the Sherpa REF service that is currently in beta testing phase. This tool is intended to be a user-friendly method of checking if a journal complies with REF open access policy
participation in the Jisc Monitor project. This project is developing tools for managing OA data. We will be testing their tools and providing data extracts direct from our system to test the aggregation and benchmarking service.

- exploring the RCUK plug-in further to assess if there are any benefits we are not yet taking advantage of. The case is very strong – RCUK say that where this tool has been used to check compliance, that will be sufficient to satisfy RCUK requirements. We are interested in seeing how this helps the institution in terms of reporting and may have more to say once we have done further quality checking based on the reports that we have set up.
• encouraging EPrints users to work together on the OA plug-in

• working on improved OA reporting for colleges and schools and sharing these with the EPrints community

• introducing more quality checks to reduce risk of non-compliance. These checks are based on the same reports used for colleges and schools and reporting to funders

• release of additional standard metadata explanations via the CASRAI UK Open Access Working Group

• further outreach to ensure authors are aware of the requirements and services. This may include encouraging more consideration of OA at submission stage as this may affect REF eligibility and there may be a cost associated with complying with funder terms.

We are sure that peer sites will agree there is considerable devil in the detail. For example:

• to address RCUK and REF requirements, we need a full publication date, where sometimes publishers only provide a year or month

• at the University of Glasgow, we process our OA costs via our library management system (Sierra) and the cost is subsequently entered onto our Finance System. This can introduce delays into the process of completing records on our EPrints system.

We are delighted to be part of a supportive OA community with an active range of workshops and discussion fora. The main methods we use for communicating about open access are the ARMA UK OA special interest group, the UKCoRR (United Kingdom Council of Research Repositories) mailing list and the EPrints UK User Group, and we try to share any developments via these routes.

Please contact us if you wish to discuss any aspect of this work.

Abbreviations and Acronyms
A list of the abbreviations and acronyms used in this and other Insights articles can be accessed here – click on the URL below and then select the ‘Abbreviations and Acronyms’ link at the top of the page it directs you to: http://www.uksg.org/publications#aa

Competing interests
The authors have declared no competing interests.

Acknowledgement
This work was supported by Jisc (grant number 3267).

References


