

End-to-End

Open Access Process Review and Improvements



Open Access and the Research Excellence Framework

Executive Summary

This report provides a summary of the demonstrations, discussions and findings of the Open Access and the Research Excellence Framework held as part of the End-to-End Project.

Presentations, supporting information and information from the lightning brainstorming session can be found on the workshop's padlet at: http://padlet.com/klw/oa_ref_workshop



Key Observations

The same areas of concern emerged as in our previous workshops, but we are finding either that solutions are being developed, e.g. the technical solutions demonstrated at the workshop, including Glasgow's Open Access workflow which is now [available](#) for EPrints, or that there are initiatives in place to put pressure on publishers and funders, e.g. [Jisc's Desiderata](#) for publishers and the Jisc Schema for Open Access Policies.

Key recommendations:

- There are still concerns about working with publishers, many of which are mentioned in [Jisc's Desiderata](#), in particular obtaining the correct metadata and getting it earlier in the publisher cycle and a need for publishers to improve their communication to authors regarding their open access options.
- Cameron Neylon recommended the use of identifiers, e.g. ORCID, DOIs, FundRef etc. and also emphasised we need to work with publishers to put pressure on submission systems vendors.

- Ben Johnson reassured people that the HEFCE OA policy was not going to change again and gave some advice on communicating HEFCE's Open Access policy (keep it simple). Focus on the benefits as well as policy requirements.
- There is still a demand to standardise metadata across systems.
- There was a request for funder policies to align (e.g. RCUK and HEFCE). In the parallel discussion, this came up as a need for a standard policy template (this is timely, as Jisc has just announced a [Schema for Open Access Policies](#)).
- We would like some guidance from HEFCE on how to self-audit in preparation for the Open Access element of the next Research Excellence Framework exercise.
- Engage with initiatives such as Publications Router and the [COAX initiative](#) which will notify all relevant institutions when a paper is deposited in a repository.

1. Outline

The [End-to-End Pathfinder project](#) ran its third workshop on 10th September 2015. This built on outcomes from our first two workshops, [Open Access: Issues and Potential Solutions](#), held in September 2014, and [Open Access and the REF: Issues and Potential Solutions](#), held in January 2015.

We had originally planned to run a repeat of the previous REF workshop, as this had been heavily oversubscribed, and was well received. However, changes to HEFCE's Open Access policy in July 2015, as well as technical developments, including the EPrints REF plug-in, Sherpa REF and Jisc Publications Router, meant that some of our outcomes and objectives had been superseded. We therefore updated and amended the agenda to reflect these changes, with a particular emphasis on the technical developments available to support HEFCE's new Open Access policy.

This workshop also proved popular and was fully subscribed, with a good mixture of stakeholders, including librarians, research managers, publishers and policy makers.

The programme focused on technical developments, with opportunities for discussion, practical demonstrations and presentations from stakeholders with different perspectives of Open Access.

2. Introduction – update on the Pathfinder Projects and the End-to-End Project

Sarah Fahmy from Jisc began the workshop with an [overview of the Jisc Open Access Good Practice projects](#), which included information about project deliverables, such as case studies, toolkits and coding in the different areas of open access, from advocacy and workflows to cost management and systems and metadata.

Valerie McCutcheon, the lead for the End-to-End project presented an [update](#) on our deliverables, which included a brief overview of the workflow [coding](#) written for EPrints. Valerie also gave a demonstration of the coding in the workshop lunchbreak. This coding has now been [shared publicly](#).

3. Lightning Brainstorming session

This was broken into several themes:

- Working with publishers

- REF Policy
- Resources
- Metadata
- Technical
- Other

Participants were asked to think about their requirements for each theme and add a possible solution if they had one. There was also a column for comments.

A transcript of the brainstorming session can be found [here](#).

The brainstorming session was then used as the basis for the afternoon's parallel discussion.

4. Key points in Implementing the Research Excellence Framework Open Access Requirements

Ben Johnson from HEFCE discussed the key points of implementing the Research Excellence Framework Open Access requirements.

Ben described the UK as having a 'settled' landscape with a mixture of green and gold open access. Open Access is a global journey; the UK is leading, but progress is being made in the EU; in fact the Netherlands is ahead of the UK and Australia is further ahead with books. The US, China and India are also engaging. There is a link between openness and quality.

The REF policy is a game changer and an achievement for Open Access.

He acknowledged that there are still technical headaches, but believes systems are ready to deposit and capture information. We still need plug-ins to capture information for compliance. The user experience is dependent on systems we have and there are cost implications if we want to take the burden off academics.

There is some nervousness around repositories and whether they are ready for this – they should be, but they may need plug-ins for compliance reporting. There are also issues around interoperability; Why should we deposit in ResearchFish and then do it all again? Work is ongoing to align funders' policies. Adam Tickell will be reporting on, and, if necessary, revising Open Access policies to government. However, Ben confirmed HEFCE will not change its policy again before the next REF, which many people found reassuring.

Publishers do not make their information clear enough – we need to place our trust in Sherpa rather than spending time looking through publisher websites. Sherpa is 95% accurate, so we should rely on it. The good news is that the majority of publishers now offer a green option compliant with the HEFCE policy.

There are cultural challenges: by far the biggest is the communication of the policy, which can be time-consuming. Keep the message simple: submit the paper to a repository or it is not eligible for REF. We also need to ensure that academics see the value of depositing by explaining the benefits better, as well as relying on the issue of compliance.

5. Working with Publishers

Cameron Neylon, Professor at Centre for Culture & Technology, Curtin University and Freelance Researcher, gave a [presentation](#) on working with publishers. This discussed the complexities for Cameron was originally invited to speak while he was working at PLoS. He stressed that most publishers are keen to help as they have a genuine interest in making customers' lives easier.

Difficulties faced by publishers:

- **New requirements.** The layout of an article is well-established but now where do the licence, embargo date and deposition rights information go? Historically they have not been needed. Now we all have different roles and interact differently - libraries are not just talking to academics in the role of reader but also now as author.
- **Communication with institution/library.** If a publisher wanted to communicate directly with library/institution then affiliation information could be found on an article but there is not an address to notify the author or his/her institution. A solution is needed so that there are structured ways of communicating, with an institution providing an authoritative source e.g. a registered end point via an API. Cameron gave an example of a perfect workflow, but this would take time.
- **Third party submission system.** In reality the publisher does not control the information taken from authors but it is run through a third party submission system. At the point of acceptance the metadata that the institution needs is not in the publishers' hands, so compliance is out of their control. Moving metadata from a submission system to publisher would be messy. Another problem is that publishers are not interested in amending their processes to suit just a few institutions or just one country, as the UK is unlikely to be the largest market.

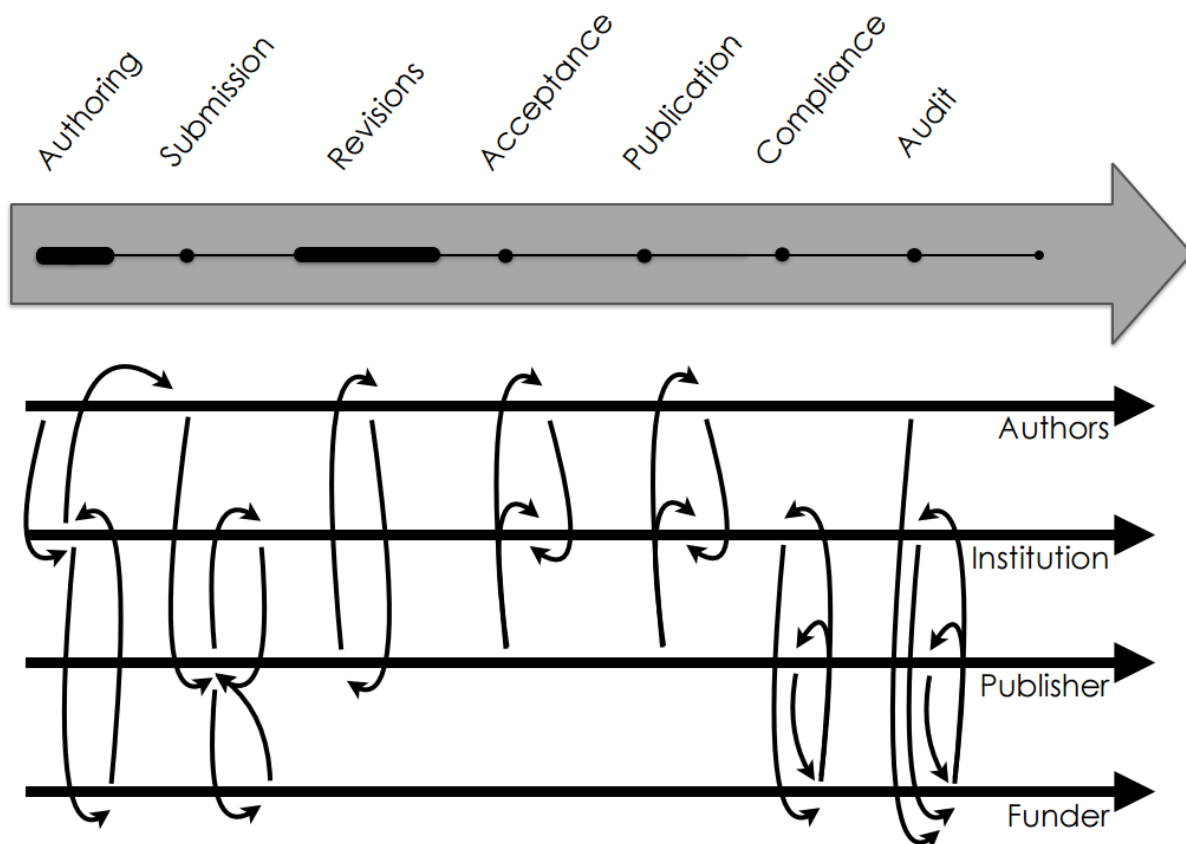


Figure 1: Neylon, C. (2015) *Working with Publishers*. Available at: <https://padlet-uploads.storage.googleapis.com/73782683/48abd5c3a8d85607d59d451cbedf8475c4240f0e/c4d9d248927397645f6c62861a66feb1.pdf>

A core message was that we need shared information models to exchange and notify throughout an article's life cycle, from early in its life cycle, before the article is submitted, and then throughout the process. A pdf and DOI are needed to be able to track an article throughout its lifecycle. Most publishers are now issuing DOI at point of acceptance so manual checks could be made on a regular basis to get the necessary metadata.

Situation in the short term:

- UK focus
- Systems sitting outside the main article flow (mainly submission systems)
- Multiple ways of routing data
- Opportunities to improve compliance if get into lifecycle early on e.g. if EPSRC article doesn't have a data sharing statement then can resolve it before it's published.

Long term situation ideally:

- Global focus
- IDs for researchers (ORCID), objects (DOIs), funders and grants (FundRef) and institutions
- Partner with publishers to put pressure on submission system vendors

Cameron was asked about community consensus, as it would need a lot of people to carry this out. A lot of Cameron's research work is about creating communities so he will start by getting individual groups to recognise there's a problem. Publishers do genuinely want to help and provide a customer service. There are many ongoing conversations in multiple locations and he sees his role as trying to steer large ships in same direction.

6. Publisher Policies, Funder Policies and the REF

Azhar Hussain then gave a talk on Publisher Policies, Funder Policies and the REF, with a demonstration of Sherpa REF. Sherpa tools include:

- Romeo - this was originally set up in 2007 to help authors. There are 2125 publishers, 26700 journals
- Juliet – includes Open Access policies from research funders, with 77 UK funders, 157 funders
- OpenDOAR - directory of Open Access repositories. 2971 repositories, 232 in UK
- FACT – shows funder compliance combining Romeo and Juliet
- REF is being developed

Open Access drivers for the tools:

- Authors wanted to know what they could deposit
- HEIs recommended that authors deposit
- Research funders required Open Access with green/gold choice; therefore Open Access became a requirement rather than desirable, which changed the game. The goal is now clear.

The Open Access landscape is changing the relationship between the author and publisher. Traditionally the author signs a contract with the publisher, but now HEIs have an interest. The

author has a responsibility to the HEI as well as to research funders and the publisher. And the HEI has a contract with the funding councils. What hasn't changed is the route to Open Access – it is still via the publisher. Contract law trumps copyright law, so should HEIs have sight of the contract?

Sherpa REF will provide advice to UK authors and provide support to Open Access advisers and repository managers. Some user journey analysis has been completed, and feedback will be sought on the prototype. It includes a feedback module and a screenshot of what you're looking at (in help and feedback). The tool focuses on the author and is building up a sequence of questions that authors may ask, e.g. which panel do I belong to? It tells you whether a journal is REF compliant and stresses that deposit must be within 3 months of publication and shows the length of embargo, e.g. Lancet would have a 6 month embargo. Where a journal is non-compliant the tool describes the HEFCE exceptions. The architecture is based on a REST-API and the Sherpa REF service is read only.

Timescale:

- prototype completed by the end of September
- user testing by the end of October
- beta launch in November

FACT and REF may be combined eventually but there's nothing to stop an institution from combining the two now.

Azhar was asked: as journal publishers propose their own licences does Sherpa REF consider licences when a report is given of compliance? No, but this licence is checked in FACT so when the two are combined it will incorporate the licence information.

7. Publication Notification – Publications Router

Steve Byford from Jisc then introduced [Publications Router](#), which acts as a go between, passing notifications of acceptance to institutions via identification information in the article to enable HEIs to capture them to their repository or CRIS.

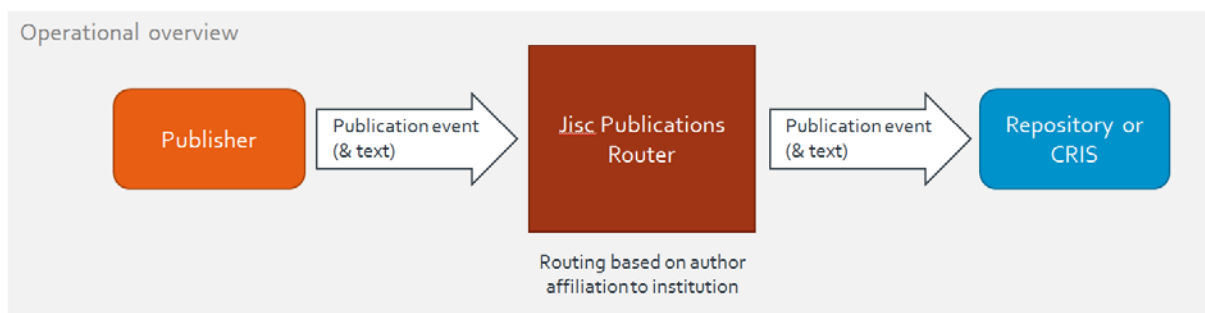


Figure 2: Byford, S. (2015) *Jisc Publications Router*. Available at: http://padlet.com/klw/oa_ref_workshop/wish/70177109

No publishers have said they do not want to be involved. However, the experience is similar to Cameron's, in that it is difficult to get good affiliation details. Publishers are not geared up to giving out information publically at the acceptance stage.

The HEI can specify parameters and decide which articles it wants. There will be no need to install an importer into the repository, but it may require some configuration.

Router mark one was developed by EDINA and demonstrated a viable prototype. It was focused on EPrints originally, but the next version will be available to various systems including CRISes. Router 2.0 is being developed by CottageLabs and will be handed over to Jisc in early 2016. It should be delivering notifications to institutions and open to new ones joining from January 2016, with plans to move it to a permanent service footing from Aug 2016.

8. Demonstration of the Open Access workflow

Over lunch, Valerie gave a demonstration of the Open Access workflow within Glasgow EPrints. The Open Access and RIOXX (labelled as RCUK for ease of use) tabs enable Glasgow to use solely EPrints for Open Access reporting. There is still some manual entering, e.g. the funding information is copied across from the finance system, and checking whether an APC has been paid is again manually checked using reports from the finance system.

In the afternoon, we had a demonstration of the new EPrints REF plug-in, as well as a parallel discussion.

9. Demo of the Research Excellence Framework EPrints Plug-in



Balviar Notay from Jisc and Tim Miles-Board of ULCC demonstrated the REF plug-in. This is not a rewrite of the REF 2014 plug-in, which covered the whole REF process. It is primarily intended to determine REF Open Access eligibility and to support audits. The plug-in will identify if an output is REF compliant, and if not give the reason. If an output is marked as non-compliant, it can then be labelled as compliant based on an exception (using a drop-down list and optional free text field).

The RIOXX2 and Dates plug-ins feed through to the Open Access tab (to be relabelled HEFCE OA), but it will work without the RIOXX plug-in.

The EPrints REF plug-in is not at beta stage yet. It is being rebadged as the REF Support Package (although this name is still subject to change).

10. Parallel session: Discussion of Other Considerations in Implementing the Research Excellence Framework Open Access Requirements.

We had a parallel session to discuss other considerations in implementing HEFCE's Open Access requirements. This was based around the lightning brainstorming exercise earlier in the day.

About half of the group worked in research support and half in the library.

Working with Publishers (this was the area we discussed the most):

- Finding out about papers. Solutions are the [COAX initiative](#) that Paul Walk is organising and Publications Router. COAX will work in a similar way to Publications Router in that everyone would be notified as someone deposits in one repository. Also Publications Router will extract information from CORE.
- We need provision of metadata earlier in lifecycle as well as when published. Suggestion that publishers should engage more with [CrossRef](#). Publishers have xml files but these are not added to CrossRef. CrossRef is asking publishers to provide more information. Improvements will be made if we can get the DOI on acceptance.
- Making licensing information more readily available with more information on publishers' webpages. Traditionally the relationship has been between the publisher and author so information is geared towards the author rather than being available for administrators. Now the relationship isn't exclusively between author/publisher but also with the institution that is paying.
- When does the embargo start? Jisc raised this question, whether it should be from the online date or hard copy publication date. The group argued that it's the first appearance of the version of record and this was the guideline to use, but we were not sure that publishers would agree e.g. publishers might have 18 months gap between the first version of record and hard copy.
- It was suggested that we used [Crossref Chronograph](#) to see when the DOI was first created and when it was first resolved. This could be used for the acceptance date.
- There is some debate over whether it is permitted to use publisher templates for accepted manuscript version. It was agreed they are usually okay to use, e.g. IEEE
- Some in the group reported that authors were confused by information provided by the publishers –author being offered the gold option but not always giving them the opportunity of green, so time and money was wasted on paying APC. Others reported that some publishers were making information about the green route more available on their websites.
- Information from publishers about licence options can be confusing and has implications for funding. For those publishers that do not offer solely CC BY licence for funded papers, time is spent asking authors to change to correct licence. Wellcome Trust won't reimburse funds if the wrong licence is applied.
- Academics are more likely to deposit in repositories if that is a requirement of their appraisal process.

Some of these issues are covered by [Jisc's Desiderata](#), which Jisc is in the process of revising.

One of issues was providing information and capturing an institutional email address to contact someone.

REF Policy:

- We felt that it wasn't exclusively the responsibility of librarians/research office to communicate policy, but Ben said it was our area to do this.
- Ben said it shouldn't just be about compliance but about benefits. There is a role in communicating how it can help research and institution as a whole.
- Good to work collaboratively - do depositing, helping with exceptions, explaining policies. Some institutions have research office which reports on REF and communicates with academics.
- Great to get the research office to work in partnership with library and put information out to academics together.
- Useful to have support from senior management.
- Academics are still unclear about the message and think it needs to be gold.
- Clarification of exceptions – although this was raised in the brainstorming session, people who attended the parallel session did not feel this was a problem.
- There was a query as to whether HEFCE would be running a 'health check' before 2017. We haven't heard of plans, but institutions are doing this. People felt it would be useful for HEFCE to provide a checklist of things to have in place, but the impression was that it was up to the institution to monitor.

Resources

- Some HEIs are employing more people to help, e.g. administrative staff to free up staff for advocacy, or for repository work.
- There was discussion of how much checking we should do. Some suggested that we should only check items being deposited in the repository and it was unrealistic not to check these records until academics can do this perfectly. Some used Symplectic as a source of reporting, although there could be problems because if the correct information wasn't input at the initial stage then there was difficulty in amending information in EPrints and getting it to talk back to Symplectic. Some were using Pure to get a basic report that was enhanced later. Others were using Pure ask for basic information that we enhance later, and on publication - may get back to author for correct version.
- Navigating publisher policies and funder policies can be difficult. Although it is difficult to change a policy (and HEFCE said they are not changing), guidance could change and a range of funders would be happy with an auditing pipeline which would set out what people need to do, rather than monitoring it.
- Specialist staff. At the moment we're expected to know everything and we are taking on additional tasks in our jobs. There are different models, e.g. research librarians who focus on areas such as Open Access and academic liaison librarians taking on Open Access within their subject roles. However, they may not know about it in same depth.
- A lot of jobs are advertised for 12 months or 18 months which are probably funded by block grants, which means they have just got to grips and then they have to leave.
- Would be useful to have a book which is an introduction to Open Access. It was noted that there had not been a book solely on Open Access since [Peter Suber's](#) in 2012, but as everything changes so quickly it would probably be out of date by the time it is published.

11. Facilitating Access to Repository Content- CORE Aggregation

Finally, Nancy Pontika, Open Access at CORE, gave a [presentation](#) on CORE Aggregation.

The mission of CORE is to aggregate all open access content distributed across different systems worldwide, enrich this content and provide access to it through a set of services. It harvests research from the UK and globally.

Currently institutions have to email CORE to have records taken down, but CORE is implementing a new dashboard facility, currently still in beta, to enable institutions to add, remove or edit content. One person per institution will be invited to administrate the dashboard, and they will be able to add others as admin.

CORE Dashboard : overview

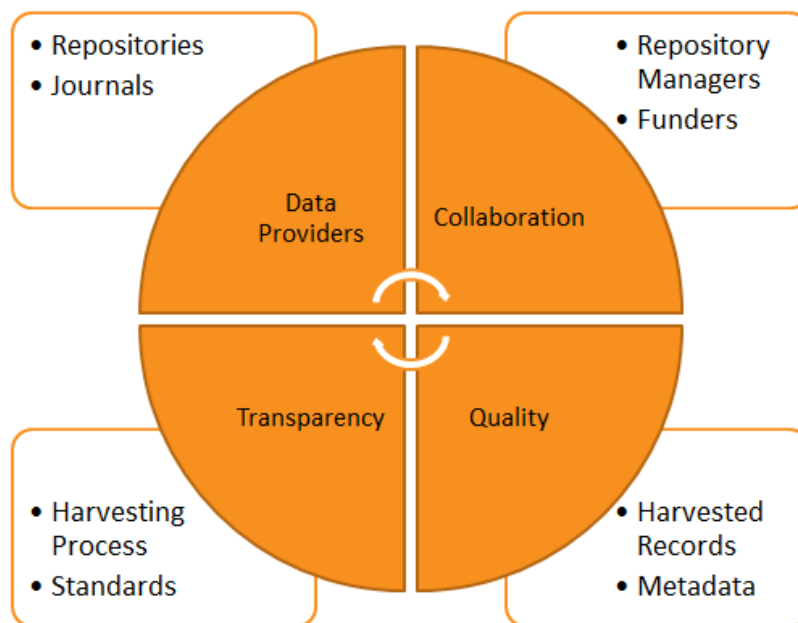


Figure 3: Pontika, N. (2015) CORE Dashboard. Available at: http://padlet.com/klw/oa_ref_workshop/wish/70177024

CORE will implement RIOXX and IRUS-UK as new functionalities. CORE will (eventually) pass records to Publications Router for co-author AMs across institutions.

12. Feedback

Our feedback was very positive; all forms rated the event either Excellent or Very Good for Content and Relevance, with 61% rating it Excellent in both categories.

Attendees liked the diversity of the programme. There was a good range of speakers, and opportunities for discussions and networking with peers at other HEIs.

People welcomed the demonstrations of different technical developments, including CORE, the EPrints REF plugin, Sherpa REF, Jisc Publications Router and Glasgow's repository workflows.

Report authors

Kate Walker, University of Southampton, Nicki Clarkson, University of Southampton, Nikki Gregory, University of Kent, Valerie McCutcheon, University of Glasgow, Michael Reilly, University of Kent, Louise Tripp, University of Lancaster

Acknowledgements

This work was supported by Jisc Open Access Pathfinder Grant 3267.

For project contact details see <http://e2eoa.org/e2e-partners>

Twitter hashtag #E2EOA or email valerie.mccutcheon@glasgow.ac.uk