

A Pedestrian View of the Roads to Open Access: Understanding Workflows to Enrich Infrastructure

Introduction

Things are moving quickly on the policy front with respect to the roads to Open Access with renewed focus on Gold Open Access¹. Gold OA is essentially the purchase of a service from a publisher: that the Publisher's (or Published) Final Copy is made available under specified terms of Open Access. This is in contrast to the Green Road to Open Access, which typically involves the deposit and availability of the Authors' Final Copy, by an author or an authorised agent.

UK RepositoryNet+ (UKRepNet) is a JISC initiative² being developed by EDINA. This is infrastructure to enable the delivery of services to universities and colleges in the UK in their work with Institutional Repositories (IRs) and the like to support Open Access (OA) and reporting arrangements for research outputs and outcomes. The focus is upon research literature as part of research output, noting the growth of the enhanced publication (i.e. the data behind the graph) and the importance of citation of data sources and instrumentation.

Though Green OA is still important, the move to Gold OA is being accelerated faster than originally anticipated at the project kick-off in October 2011. Accordingly UKRepNet was given a 'watching brief' to monitor the significance of Gold OA in order that infrastructure being created remained relevant.³ What follows is report from that Watch activity, setting out the envisaged workflows, both financial and informational. The RepNet focus is on delivering a sustainable and fit for purpose repository service infrastructure. This has an emphasis on institutional repository infrastructure and therefore a focus on Green but naturally there are links to Gold, and as other service requirements emerge the service environment would need to be extensible to encompass those or to at least interoperate with them. Related to this JISC initiative & the wider OA environment (both Gold & Green) is the work on metadata to research outputs that JISC is taking forward with publishers & RCUK: this includes RIOX, Vocabulary for OA (V4OA) & also the activity via the Open Access Implementation Group (OAIG) on APCs (which this short report is feeding into).

RIOX & V4OA constitute action with regards to the agreement on UK Repository Application Profile to include OA and Funder Metadata. The semantics of OA are being dealt with in V4OA which is a JISC action as a result resulting from the agreement to address this issue between OAIG & Publishers. This work will feed into the National Information Standards Organisation (NISO).

¹ "Accessibility, sustainability, excellence: how to expand access to research publications: Report of the Working Group on Expanding Access to Published Research Findings", accessed on Jul 18th, <http://www.researchinfonet.org/wp-content/uploads/2012/06/Finch-Group-report-FINAL-VERSION.pdf>
EC Communication COM(2012) 401 final, "Towards better access to scientific information: Boosting the benefits of public investments in research", http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf

² UK RepositoryNet+ Project, <http://www.repositorynet.ac.uk/project-summary/>

³ Component services in UKRepNet will include RoMEO, JULIET, R_J Broker and IRUS-UK.

Understanding Workflows to Support Open Access – and Research Information Management

This builds on work by the JISC-funded SONEX Workgroup on deposit opportunities into repositories and motivation for interoperability, using a multi-authored work from more than one institution⁴.

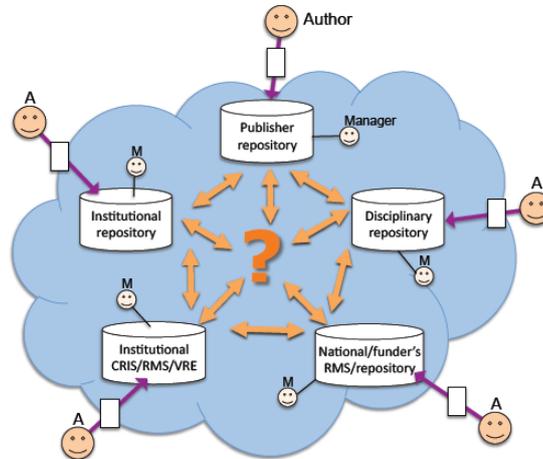


Figure 1.- Deposit usecases: Information exchange among diverse repositories

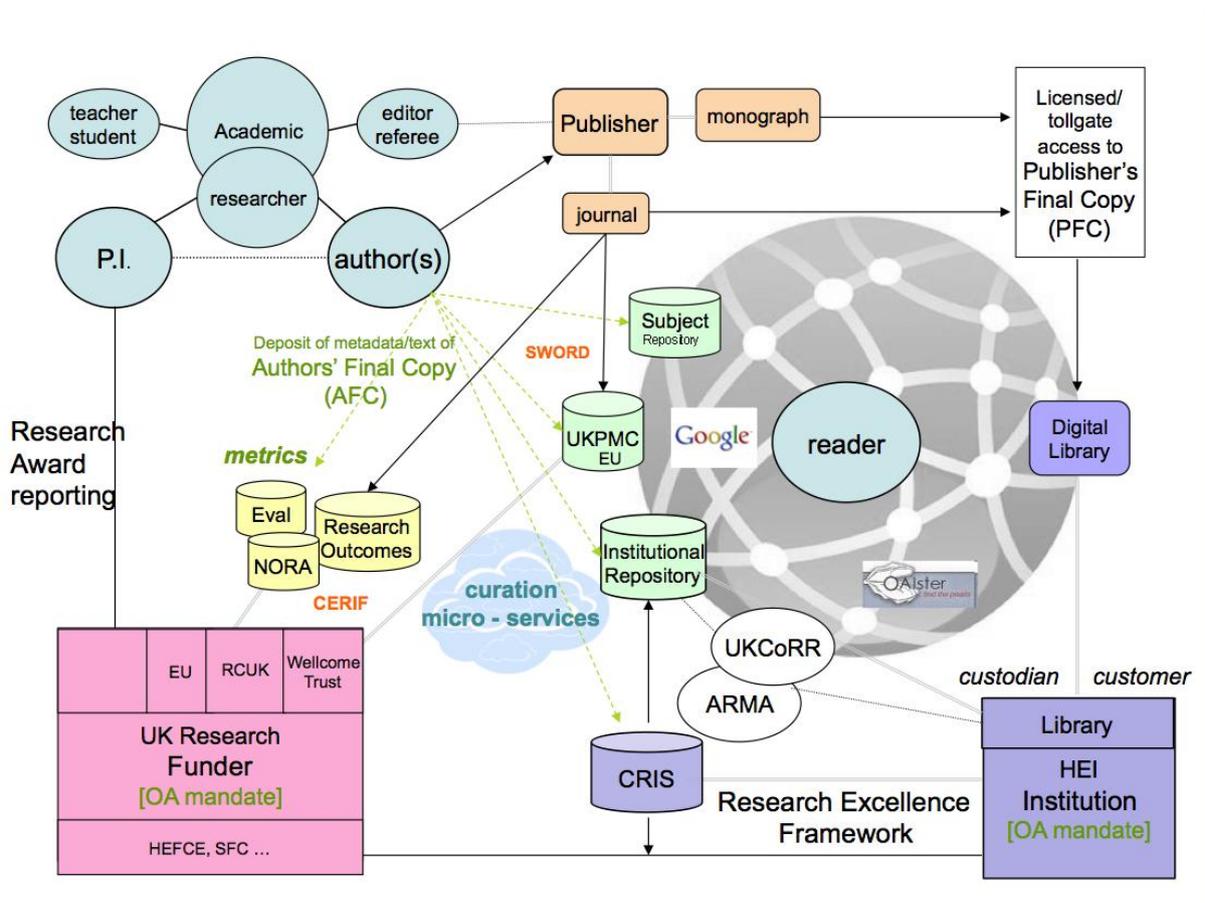
During the latter part of 2011, the UKRepNet Team set out a diagrammatic representation of the Repository Landscape which identifies a number of Actors: Funder, Institution, Principal Investigator (PI), Author and Publisher. Shown below as Figure 2, this includes a number of bilateral workflows between those Actors, both for the purpose of publication and dissemination of research outcomes and also their reporting. What needs to be assessed is the extent to which this can generate understanding for both Gold and Green Open Access. In either case, these workflows are firmly based within the perspective of the Institution in which the PI plays a leading role in decision making with fellow researchers on authorship and choice of mode of publication, including the choice of journal. What is less clear is the subsequent reporting to Funders.

The story begins, top left in Figure 2, with research proposal and award of grant that gives obligation of report by the Principal Investigator (PI) to the Funder. Then there is outcome of research, with submission of the (multi-)authors' manuscript to a journal, part of what might be regarded as the flow of publication from funded research⁵.

⁴ Figure 1 is taken from Burnhill, P, Castro, Pablo de, Downing, J, Jones, R, Sandfær, M, "Handling Repository-Related Interoperability Issues : the SONEX Workgroup", <http://hdl.handle.net/10016/9257>

⁵ See page 8 of Burnhill, P, & Tubby-Hille, M (1994). 'On measuring the relation between social science research activity and research publication', *Research Evaluation*, 4(3), pp130-152. Text available online as eig.sdss.ac.uk/projects/rapid.pdf

Fig. 2 Actors, agency & relationships for report, deposit & access



Institutions and Funders are two key stakeholder groups, with variety of motive, noting that all institutions are not of one type: ranging from the large research intensive universities to less well resourced small to medium sized institutions. Each of these two key stakeholders generates workflow and controls elements of metadata needed by the other. The Research Excellence Framework (REF) is noted for its importance for the institution, alongside its need to satisfy compliance with the requirements of Open Access mandates by Funders and Institutions. This is driving other workflows.

One notable characteristic is the deficiency in metadata for those workflows, typically that only one of the two key reference identifiers is present: the Funder's Award Reference Number and the Digital Object Identifier (DOI). Publishers wish the Authors' Final Copy to contain DOI link to the Publisher's Final Copy (PFC) and the citation to this published version is also wanted by the authors for purposes of impact. These issues are being dealt with via the metadata initiatives referred to above.

Modelling Workflows Along The Green Road

Green OA as such does not require a payment mechanism; the focus instead is on making the ‘Open Access representation’ of the published work openly available. The motivation for Green OA derives from policies adopted within the academic community, including mandates by institutions and by funders, with the intention of increasing timeliness, visibility and impact. This is best exemplified by ArXiv, the open access repository hosting eprints of research articles on physics, math, non-linear science, computer science, quantitative biology, quantitative finance and statistics hosted by Cornell University.

Green OA generally involves the deposit of the Authors’ Final Copy (AFC) into repositories, both institutional and subject, shown as green dotted line in Figure 2 and as indicated in Figure 1, above. The Publisher also has the Authors’ Final Copy, and at least one major publisher has been working in a pilot study to assist the Author in complying with Funder mandates.

What is deposited by the author is not always the AFC: for example some funders (eg Wellcome, European Research Council), mandate deposit of the ‘version of record’ into UKPMC .

Information workflows associated with Green OA need to account for the ‘embargo’ limitations placed by publishers on what can be done with the Authors’ Final Copy, and when it can be made available under OA. (RJ Broker, which will be managed by RepNet, has functionality to manage embargos).

Modelling Workflows Along The Gold Road

The purpose of this paper, as part of the ‘watching brief’, is to understand how to gain leverage from an informational point of view from the variety of initiatives being proposed for financing Gold OA. This is especially for the inclusion of the two essential pieces of metadata: the Award Reference Number and the DOI for the PFC⁶ and to ensure we can be extensible, interoperable and future proof.

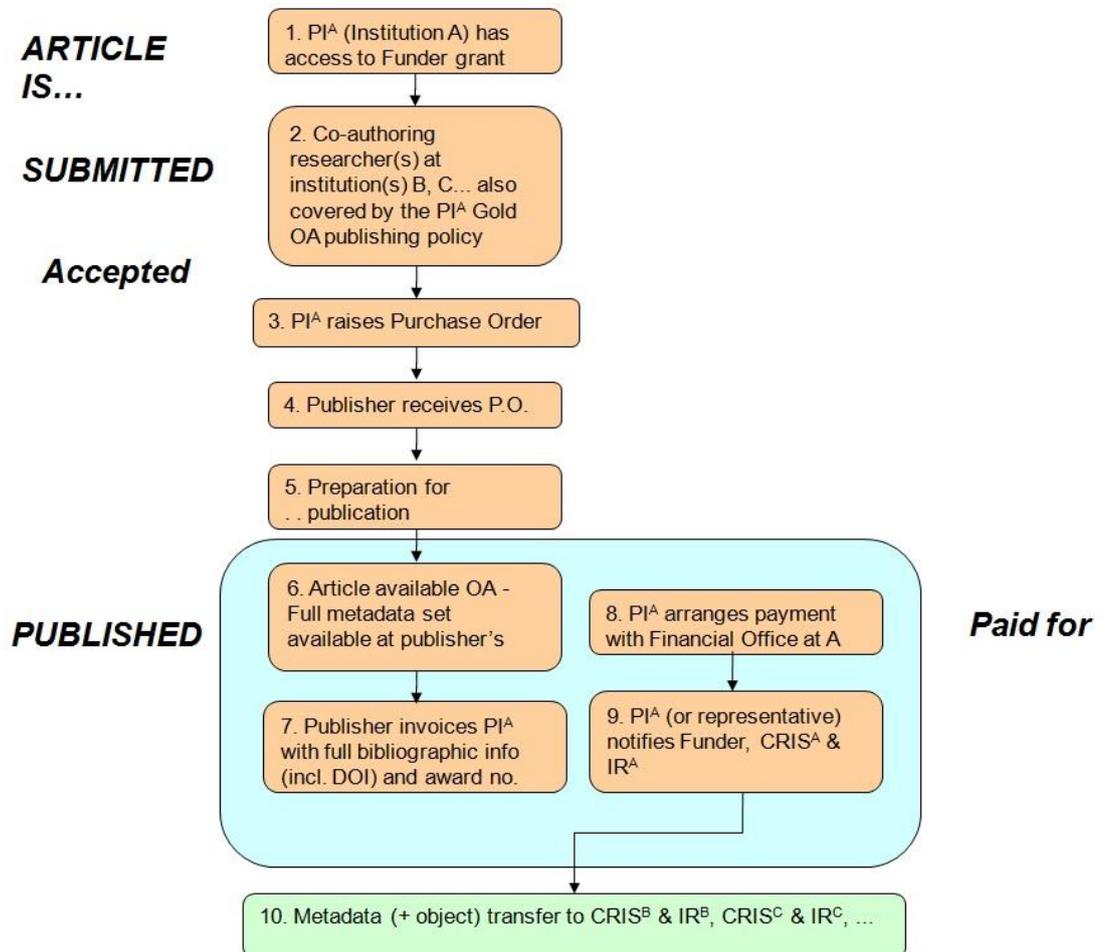
The Finch Report is not understood to imply a reduced role for Green OA or for Institutional Repositories (IRs) in the scholarly communications process⁷. The researcher, both as author and as a Principal Investigator responsible within a university for decision-making about a research grant, remains a key actor. Other actors include academic support staff within the walkways of such an institution, and of course both the Funder and the Publisher.

The Gold OA workflow has its focus on the Publisher’s Final Copy. It includes the same range of actors but there may be greater interaction between the PI for a funded project and the Author. The following is a simple representation of how such a financial workflow might work within a Gold OA landscape, shown both graphically and as narrative.

⁶ Both of which should be part of the guidance provided by the RIOXX Project ‘RIO Extension: Mapping Repository Metadata Requirements’ in which UKRepNet is involved, http://www.jisc.ac.uk/whatwedo/programmes/di_researchmanagement/repositories/rioextension.aspx

⁷ A measured response to Finch, representing the views of academic librarians through SCOUNL and RLUK and addressing the transition period to full Gold OA, was issued on 25th July 2012 - RLUK/SCOUNL response to BIS statement on access to publicly funded research, <http://www.rluk.ac.uk/content/rlukscounl-response-bis-statement-access-publicly-funded-research>

Fig. 3 Modelling the transactional flow for Gold OA



1. The Principal Investigator (PI) in Institution A is awarded a grant for a research proposal from a Funder, having Funder Reference Number XXX, in which there is agreement that there may be use of funds (Gold) to make research publications available under terms of OA – typically using CC-BY.
2. A Researcher in Institution B working in the multi-institutional project led by that PI (in Institution A) is in correspondence as an Author with Publisher who offers an OA (CC-BY) service for a given article that has been submitted and has been accepted for publication, in a Serial having an ISSN. Typically this article has several authors, from several different organisations. The Author and Publisher complete the submission/acceptance process with regard to publication of article in said Serial.
3. On article acceptance, the Author communicates with PI who having agreed that this is an appropriate use of funds in the grant raises a Purchase Order (PO) within Institution A. That PO specifies the service that is to be purchased, including the bibliographic detail, the terms

of OA required (eg CC-BY) and the Funder Reference Number. (Typically only the details of the Serial are known at this point (eg Title and ISSN) as those for the article (authors, title, issue/volume) may still be a bit hazy)

4. The Publisher receives PO (with the specification explicit, including the Funder Reference Number), against which they would subsequently be able to raise an Invoice.
5. Once the Publisher has arranged publication in an issue of the Serial the full bibliographic information for the article is determined, including the assignment of the DOI; there is also opportunity to include Funder Reference Number in the article and its metadata.
6. The Publisher is then in a position to make the article (with DOI) available under the specified terms of OA, according to the PO.
7. The Publisher, having carried out the service that was purchased can then present the Invoice to the PI (copying in the Author) for payment. Both the Funder Reference Number pointer to the full bibliographic information (including DOI) would allow delivery of the service to be checked.
8. The PI arranges payment of funds from grant (Funder Reference Number) with the Finance Office at Institution A: the PI is assured this will follow when there is a Purchase Order against which to match an Invoice.
9. The PI and the Institution (via CRIS, IR or some system) can now notify the Funder, supplying all required information including Funder Reference Number and the DOI, with pointer to the full bibliographic information.
10. The IR/CRIS at Institution B, C etc are notified that the article has been deposited as per mandate and is now available under OA, under, for example, CC-BY.

Note that this workflow could make use of the CERIF record structure.

The modelling above treats the purchase of Gold Open Access much like any purchase using funds managed at the Institution with delegated authority from the Funder. There would need to be additional agreement between the Institution and the Funder to operate this beyond the formal award period.

The workflow needs little modification in order to work with the 'publication funds' that universities are invited to establish with grants from the seven UK Research Councils in order that they could pay publishers an 'article processing charge' (APC) to publish their work.⁸

Whatever new agents⁹ may emerge eager to assist publishers, this should not be allowed to have a negative effect on the flow of metadata in systems required by Institutions and Funders.

⁸ <http://www.out-law.com/en/articles/2012/july/government-details-open-access-funding-and-publishing-policy-for-taxpayer-funded-research/>

⁹ For example, Open Access Key (OAK), <http://www.openaccesskey.com/>

UKRepNet and OA Mandates: Green or Gold

Universities and Science Minister David Willetts has stated in his formal response to the Finch Committee that "[each] Research Council... will monitor compliance with its policies at grant level through its outputs systems. Once Research Councils have established the payment mechanism, operational details will be set in discussions with the academic community."

The Dashboard below is conceived as a tool for monitoring funder mandate compliance rates in a way that is agnostic with respect to Green or Gold. It can be regarded as extending existing features, such as the UKPMC Grant Reporting Service¹⁰, to the whole UK repository network.

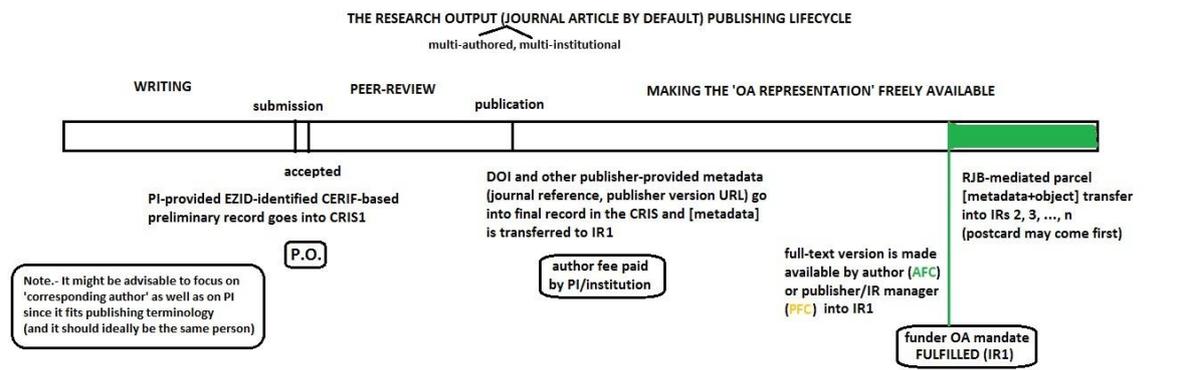
Fig. 4 RepNet data-driven infrastructure use case 1 – The Dashboard

RepNet Use Case 1: Monitoring Funder OA Mandate Compliance Rates - THE DASHBOARD

Note 1.- This approach is OA-model-agnostic: it's about compliance, and whether it's via Green or Gold OA it's set by the publisher policies (RoMEO)

Note 2.- The global picture becomes clearer when taking a two-sided approach: the *finance side* on the one hand, the "data-driven infrastructure" on the other

Note 3.- The "principal deposit" concept is critical: default deposit into the CRIS/IR compound is done by the corresponding author while RJB will deal with the rest



The Dashboard is an example of work in cooperation with the Project Innovation Zone at UKOLN on the analysis of a set of use cases intended to realise the data-driven infrastructure approach¹¹, and as part of gap analysis to assess the need for additional components and services for the UK repository network.

¹⁰ Alison Henning, 'My UKPMC': new grant reporting services, <http://bit.ly/NoW1sn>

¹¹ JISC Observatory draft TechWatch report "Preparing for Data-driven Infrastructure", <http://blog.observatory.jisc.ac.uk/techwatch-reports/data-driven-infrastructure/>

The workflow could feature in any plans to build infrastructure to support mandates for Open Access, including the variety of initiatives known to be currently underway on processes for managing Gold OA payments in Europe¹².

Peter Burnhill, Pablo de Castro, Andrew Dorward and Terry Sloan

UK RepositoryNet+, EDINA, University of Edinburgh

August 2012

¹² In Germany, 18 universities receiving grants from the Deutsche Forschungs Gemeinschaft (DFG) for Gold OA Publication funds have worked out different workflows to address this. We are monitoring the COAR project looking at workflows and payments for Gold OA.