Open Access Technical Workshop
'Un-Report'

4th April 2016, Glasgow

This 'un-report' provides a summary of the Open Access (OA) Technical workshop held jointly by the End-to-End and LOCH Projects. The workshop was highly interactive and feedback received indicated that it was extremely valuable and provided opportunity for a useful exchange of ideas.

Presentation and workshop slides are available from:

(blog post is ready to release once report is ready to go on there

Report Authors:

Valerie McCutcheon University of Glasgow valerie.mccutcheon@glasgow.ac.uk
Dominic Tate, University of Edinburgh Dominic.Tate@ed.ac.uk

Key contributors are acknowledged in the report. We also acknowledge the contribution of all who attended.

Key observations:

- Scottish Open Access Group to be formed. The authors will take this forward.
- System agnostic definitions and specifications for metadata are important to aid interoperability and facilitate transfer of information between research organisations and report to funders.
- Need discussion with publishers to get metadata available for import sooner
- Confusion over HEFCE REF and RIOXX differences needs further conversation
- The policy decision to allow OA compliance within 3 months of publication date is causing some concern at institutions. Most are sticking with the message that action should be taken at acceptance stage to avoid causing confusion and as that is a good time to have access to the final agreed manuscript and provide support to users.
- Need SHERPA REF to be amended to facilitate the acceptance stage policies rather than the bottom line of publication date as current offering cannot be utilised to its full potential without causing confusion and possible error.
- Concern over how to resource compliance with all of these open access requirements.

Index	

Overview

The workshop, held on 4th April 2016 brought together a group of around 50 stakeholders in administration of open access systems and processes. Most attendees were from research organisations.

The workshops aimed to help share best practice in managing open access within systems with particular focus on Research Excellence Framework (REF) requirements.

There were short overviews of 7 different systems that had some sort of functionality to deal with open access requirements. Figure 1 shows a summary of some of the relevant functionality.

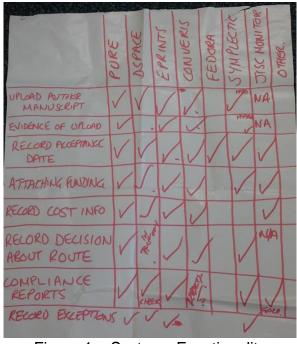Notes below are informal and sometimes verbatim.

Figure 1 – Systems Functionality

Presentations on Repository and Research Information Systems

1. PURE
Manya Buchan, Elsevier
m.buchan@elsevier.com; @manyabuchan)

- Recent focus on OA functionality -compliance for REF
- Publication status and reminder that full date of acceptance required
- SHERPA ROMEO info shown after inputting journal title
- Electronic version - add file - properties include version
- Can capture Article Processing Charge (APC) information
- Capture relations including grant (Project)
- Can save document (evidence) against acceptance date
- History & Comments for decisions
- Save and look at assessment section (REF compliance status)
- Note for personal users to beware of changing anything
- No deposit within 3 months shows as non-compliant, unless exception added
- Monitoring screens for tracking
- Reporting also available with all relevant dates, versions and status

(See Appendix 1 for notes from the PURE breakout session)

2. DSpace
Claire Knowles, University of Edinburgh
claire.knowles@ed.ac.uk; @cgknowles

- New REF form for DSpace users without a Current Research Information System (CRIS) – developed on collaborative basis including working with Jisc
- Lots of dates on the form – metadata fields under the hood which capture the data
- Currently a temp solution - @tmire building a plugin which will include monitoring dashboard
- Combined input from Pure, EPrints solutions - EPrints and DSpace to provide feedback and input to JISC on future requirements, short comings and next steps for REF plugins.
- @tmire plugin will include compliance checks. Due for testing 4 May – 6 Jun
- Input will include ORCID lookup
- Form includes licence URLs, dates, versions and exceptions, as well as adding projects
- Final tab will show green/red ticks to show compliance
- Can add fields fairly easily. Reporting on these would be additional work

(See Appendix 2 for notes from the DSpace breakout session)

3. EPrints
Valerie McCutcheon, University of Glasgow
Valerie.McCutcheon@glasgow.ac.uk; @mccutchv

- Customisation done in Glasgow, available to the community
- Upload and evidence acceptance – acceptance emails saved
- Attach multiple dates and funding
- OA tab for recording APC information – avoids separate spreadsheet
- Tag as 'gold', 'no green option' etc.
- REF tab to record exceptions
- Reporting - Glasgow has produced and shared

- Still to - automate embargo release, link with REF selections module, improve reports, work on standard vocabulary (CASRAI)
- Decided to store acceptance emails, still to decide if will continue
- Compliance checking somewhat a manual process
- May refine some fields in consultation with wider EPrints and Open Access communities
- Questions:
    - Q: What currently happens with release of embargoed papers?
- This is one of the things where we have manual intervention to check which items have moved to published stage. We are currently working on improving this if anyone wants to work with us or share their solutions for this.
- Q: How much technical support is needed to install the customisations for compliance/OA?
- Very little. Glasgow are happy to help with this – contact us valerie.mccutcheon@glasgow.ac.uk
- Q: Can you add RIOXX-compliant funder ID numbers?
- Our version of EPrints if fully RIOXX compliant and we import official funder references from our research system. We are also exploring a standard funder module with EPrints peers. Please get in touch if you want to know more.

4. Converis
Kirsty Wallis, University of Kent
k.r.wallis@kent.ac.uk; @OsMonkey

- Still in testing/troubleshooting stage with Publications module
- In the test version acceptance date can be added and funding can be attached. APC purchase and related fields added
- Reporting in development – no compliance checking available yet – so currently using EPrints with RCUK (RIOXX) and REF plugins
- Hint text with acceptance dates well received
- APC set as a question – who paid? Useful if Library not involved in Gold
- System will talk back to Project module and display funding and funder details that way
- Concerns about eventual integration, dependency on international vendor for UK-specific requirements, potential Higher Education Funding Council for England (HEFCE) review

5. Fedora and Hydra
Masud Khokhar, University of Lancaster
masud.khokhar@lancaster.ac.uk; @mkhokar

- Using Pure for OA. Exports publications to EPrints
- Investigated uncommon systems (Fedora/Hydra)
- Pure is simple and familiar. Main concern is lack of good API and web service options, and not fully RCUK (RIOXX) compliant. OA compliance reporting not yet easy or useful. New fields take a long time to implement through User group requests
- Fedora based on Resource Description Framework (RDF) and linked data. An architecture, not a ready-made repository. Agnostic to content or metadata so can be a single code base for all needed features
- Easy to map RCUK(RIOXX) in to Fedora and add fields if necessary
- RDF not designed for nested structures, difficulty with e.g. author attributes in a single element
- Must have technical people. A blank canvas but steep learning curve and complex mapping.

- Hydra – based on Fedora – designed to simplify Fedora as a ready-made solution though still needs customisation. Can use Sufia to combine required components for an OA repository. Upload management and version management really good – can select the version you want to display from multiple options
- Issues: Needs custom development. No OA reporting out of the box
- RCUK (RIOXX) playground – form with full capture, to push into Fedora or Hydra. Not yet possible with Pure as lacking full RCUK (RIOXX) compliance
- Have stopped requesting acceptance date evidence – improving process efficiency
- Questions:
  - o Q: Are you planning any CRIS integration or stand-alone?
  - o Q: Controlled Vocabs – silos? How are data exposed?

6. Symplectic

Nicola Barnett, University of Leeds/Leeds Trinity University

n.s.barnett@leeds.ac.uk; @NicolaBarnett74

- Deposit encouraged at acceptance. Assumed deposit done by academic – go to 'Deposit your work' link
- Select type, can select with Title or DOI - or skip and enter minimal data manually. If status of 'Accepted' selected, accepted date will become mandatory. (Likewise 'published' and date of publication fields). Library staff check all data and flesh out the record
- Select pre-populated grant or 'not externally funded'
- Upload file OR enter OA location – people tend to add things like academia.edu so use of OA location is not encouraged
- File version is mandatory field
- Library ALWAYS check embargo dates where specified by the author
- File deposit date is recorded automatically and is visible
- OA Monitor module added on – pick School/Faculty and will show anything within scope of REF policy, with various filters.
- Compliance filter will allow you to select specific REASONS e.g. no deposit, too late, excessive embargo length, no exception
- Deposit filter – will pick up anything without file that will go past deposit deadline, choice of timescales, e.g. due in 14 days, overdue
- Problems – will pick up anything without dates so includes old items – need filter to EXCLUDE nil dates. Looks too closely at embargo lengths and includes items without embargo set (not yet published)

- Q: Not all conference papers are in scope for REF policy - only those with ISSNs. Are they all measured?
- A: All conference papers are treated as if they are in scope.
- Q: How possible is it to customise metadata fields?
- A: It appears that they can be customised fairly easily. We have added fields in advance of them being added in later releases by Symplectic. E.g. we added our own date of acceptance field before Symplectic released the compound date fields.
- Q: Do you check the right REF eligible version (not just if deposited first version)?
- A: Yes, we check files when we receive them to make sure they are the AAM, or VoR if gold.

Figure 2 – Presentation on Symplectic

7. Jisc Monitor

Valerie McCutcheon, University of Glasgow
Valerie.McCutcheon@glasgow.ac.uk; @mccutchv

- Involved in Jisc Monitor pilot – not yet live
- Monitor Local:
  - Is not a complete repository tool, focus on APC management
  - AAM deposit out of scope
  - Acceptance date and Funding info included
  - Recording Cost info – details against specific articles
  - Includes funder acknowledgement and research materials statement (data) for RCUK compliance
  - Reports to check compliance – being tested

- Monitor UK: Data supplied and aggregated at UK level for benchmarking etc.
-
  - Q: Will there be an API to feed data from HEI to Monitor UK?
  - A Yes
  - Q: Is it a standalone thing, or a plug in to repository software? (Which software? Or does it link up?)
  - A Standalone. Monitor Local is an application that can be used to manage APC data. Monitor UK is more of an aggregation initiative to take data regardless of system and use it for negotiation and benchmarking.

Generic Questions and Clarification on Systems Session

- Adequate help tips for the user essential
- How much usability testing is being done – for end users not just admin users? Date fields particularly can be very difficult for authors. Should systems have SAME guidelines to make sure aggregated data makes sense?
- Need discussion with publishers to get metadata available for import sooner
- Protocols on metadata mapping needed to allow better interoperability. E.g. could not import from Router to PURE.

- HEFCE and RIOXX inconsistent – clarification needed
- Need generic ability to export via Open Archives Initiative Protocol for Metadata Harvesting (OAI-pMH) – still an issue with different implementations

Break Out – Non-System Requirements

- The RCUK budget for 2016-17 has not been announced yet. It would be useful to have more clarity about this, both for the larger universities which have previously been in receipt of larger grants, and for those institutions which have not previously received a grant.
- Some institutions have reported difficulties fully spending the RCUK block grant. This is normally because of lack of awareness amongst faculty, money for APCs coming from elsewhere or an institutional emphasis on the green route.
- If gold OA is perceived as an easier way to get REF compliance then we may see an increase in uptake of RCUK OA funding as academics perceive the admin process to be easier.
- Gold OA can be problematic for un-funded academics that want to publish in fully gold science journals but do not have any available cash for APCs. Some institutions have small pots in place to cover these APCs.
- What about the relationship between RCUK grant and participating in offsetting deals e.g. Springer Compact?
- Is there a freeze on COAF (Charity Open Access Fund) payments?
- Acceptance date – feeds from publishers not possible/proving difficult to access data.
- Issues around ORCID and acceptance date.
- Sherpa REF -at the moment aligns with current FC policy – date of publication – however HEI are telling faculty date of acceptance – in line with the policy as it will be in one year. Server services are aware of this and our wording. We are waiting for more clarity around exceptions/tolerance of gold.
- Resourcing and providing adequate staff to manage the publications requirements for REF is clearly a concern. Staffing e.g. contact time re-advocacy and processing.
- What levels of mediated deposit are out there? Mixed, sometimes dependent on different schools, but sometimes for individuals. Self-deposit at Queens. Responsibility lies with the researcher. Glasgow School of Art research committees don't see outputs unless they are put in the repository. University of Central Lancashire some schools have appointed a person to deposit on behalf of faculty. Difficulties of dealing with data from new staff is that may have no relation to HEI or full REF.
- Resourcing issues are certainly a big concern.
- Discussion around the date of publication versus date of acceptance different HEI is giving different messages somehow very firm guidelines I think if all of the information and recommend an approach.
- Are there too many systems? There are probably too many priorities/standards. This breeds inertia. Academics overwhelmed by systems too (repository, research fish, academia.edu, subject repository).
- External pressures separate from institutional policies.
- Which systems support RIOXX? E-Prints, DSpace & Converis (in pilot).
- Is RIOXX required for REF?
- Jisc Publications Router will be compatible with all systems but testing and development work is needed. We need APIs to participate.
- REF compliance evidence gathering: Some HEIs store emails etc. some do not.

- Recommendation to keep as much as possible for audit and keep the date if you have it (suggested by people who have been through REF audit before). Should we be going back and asking for date of acceptance evidence respectively? Do we need this? If yes or possibly, then quite possibly we should.
- What about late-stage PhD is, postdocs, ECRs, who don't know whether or not they will be submitted to REF? (Author's note – we would treat them all as potential eligible so take the same compliance steps)

Other Tools for Delivering Open Access

CASRAI
Valerie McCutcheon, University of Glasgow
    uk@casrai.org; http://dictionary.casrai.org/Main_Page

- Started in Canada, now a UK chapter
- Key benefits – shared protocols for better efficiency – commonality
- System agnostic. Topics create working groups, can build on previous international work
- Lots of strands ongoing. Can join approx. £350/year. Initial meeting May
- Members could join steering committee, propose working group. Anyone from HEIs able to review and comment
- OA work – avoid reinvention.
- Can then revisit and improve any definitions produced as required.
- Technical interoperability – could choose common data model e.g. CERIF to deliver requirements
- Idea: Award letters – could data be imported directly to HEI systems

Sherpa REF
Balviar Notay, Jisc
    azhar.hussain@jisc.ac.uk;

- Sherpa/REF fits into suite of services at Submission
- HEFCE funded beta service – aims to meet OA requirements
- Aimed at authors and research support staff "I'm about to publish, does my journal allow me to comply?"
- Aware of discussion on UKCoRR around deposit requirements and wording – please feedback via the buttons on the site or to azhar.hussain@jisc.ac.uk
- Concerns:
    - HEIs would prefer acceptance message not deposit at Publication - May not be able to codify HEI policies – though in future perhaps could use Juliet?
    - Difficult with showing Gold OA as compliant when HEFCE have it as an exception
    - Is anyone recommending Subject Repositories? If not, why would Sherpa/REF emphasize this option
    - UCL Pathfinder mentions use of subject repositories https://blogs.ucl.ac.uk/open-access/2016/02/04/workshop-summary-ref-exceptions-subject-repositories/
    - Any plans for an API? – Yes, beta already available

Jisc Publications Router
Hannah DeGroff , Jisc
    https://pubrouter.jisc.ac.uk/

- Fits in to lifecycle at Deposit/Publication
- To capture articles and metadata, and alert HEIs
- Europe PMC and eLife in place, 2 new publishers to be announced at UKSG
- Further publishers in negotiation
- Key challenge – scalability – capturing UK outputs within global context
- Will source data from PubMed and Crossref
- Publisher systems just not set up to provide data easily
- Workflow: data via ftp pushed (can be harvested via OAI-PMH) to a system that needs to be RIOXX compatible
- De-duplication needs to happen in cRIS/IR
- First batch of HEIs will ingest into EPrints. Talking to cRIS systems – might need mapping work
- See https://pubrouter.jisc.ac.uk/ to read more, sign agreements and get started
- First step will be upload csv file with matching parameters e.g. HEI name
- Notifications will show in the Router, and the same ones would be available from within EPrints
- Contracts required to ensure embargo terms will be respected
- Questions:
    - Q: assuming some systems won't have SWORD or OAI harvester; could user community develop some way to ingest?
    - A: Yes, could configure API and do some mapping

Funding Acknowledgements

Open Access Technical Workshop

E2EOA and LOCH Pathfinder Projects, Glasgow 04 April 2016

Pure breakout session

Manya's demo was using v5.6, due for release in June. The features demonstrated are available from v5.5.0, and the user role demonstrated was Administrator.

There are various configuration options available to limit the visibility of certain REF-specific fields (e.g. the REF OA Compliance Status), where desired. The REF OA Compliance monitoring screens are visible to all users with access to REF2 in the REF module.

Questions on functionality

Q: Can REF OA Compliance summary screens be exported for sending to Directors of Research?
A: Can't export but can share by cutting and pasting into Excel

Q: Can we add new metadata fields?
A: Can rename existing fields but shouldn't repurpose as would cause migration or interoperability issues. It is recommended that new fields are requested via the User Group. REF-specific fields are more easily deliverable as they are UK-specific and therefore do not affect other customer bases.

Q: Which version has the new reports as shown, and how do we create them?
A: v5.5.0 has improvements, and some reporting templates are available in the Release Notes. Manya will also share them via REF wiki. They include dates and publication states in separate fields. We might need to configure our reporting module to include non-standard data fields

Q: What is needed to make Pure RIOXX compliant? Is it on Roadmap?
A: The facility to record multiple licenses for a single electronic version, with individual start/end dates, is not available. Additionally, the Group noted that compliance with the Free-to-read element is required in order to fulfil RCUK reporting requirements (funder ID reference reportable against publications). Manya noted the use case for reporting to RCUK and will take this back to the development team.

Q: Does Pure provide OAI-PMH?
A: Currently doesn't expose everything and doesn't have any functionality to harvest. More help and documentation (including examples) was requested, as there only seem to be partial results in CORE

Q: Can Pure create and store 'text mineable' versions of publications?
A: Pure cannot create 'text mineable' versions of publications, but can store them and make them publicly accessible.

Q: Can Pure record whether an uploaded file is 'text mineable'?
A: Currently doesn't offer functionality to record whether an uploaded file is 'text mineable'. Manya to investigate whether it is possible for Pure to automatically determine whether an uploaded file is 'text mineable'.

Q: Can individual institutions change to 3 months from publication date for the 1st year of REF?
A: As agreed with the User Group, Pure currently monitors REF OA compliance against the acceptance date. Once the final policy has been published, we will make the necessary changes in Pure.

Q: Does Pure support automatic embargo release?
A: Yes. Once the embargo date has passed, the electronic version is made available automatically.

Q: Does Pure capture ORCIDs?
A: Yes, and supports the export of content to ORCID.

Q: If no acceptance date, does item appear in REF OA compliance monitoring?
A: No. To be included in REF OA compliance monitoring, the full DD-MM-YY acceptance date must be input. Customers can use reporting to identify those outputs missing the full acceptance date.

REF compliance concerns
- Aware that Pure's REF OA compliance functionality does not deal well with fluctuating properties. Sometimes need to save or revalidate, but even then doesn't always update e.g. if Open item reset to Embargoed
- Further work is required on the REF OA real-time tracking screen to ensure that multiple electronic versions with varying properties are reported on accurately.
- As REF OA compliance is currently being measured against the acceptance date, rather than the publication date (which is contrary to the policy, which for at least this first year, will measure compliance against the publication date), some outputs will be flagged erroneously as not compliant.

Tips and discussions
- Attendees liked the more workflow-based approach to deposit for authors provided by some other systems, encouraging authors to add only essential data. Lancaster have highlighted 'mandatory' fields (i.e. required but not functionally mandated) in their templates in Pure and will share how this is done
- Would prefer date of acceptance to be a mandatory field to capture full DD/MM/YYYY, but as dates are shared across other publication states, this can't be done. Would be better to make Accepted/In Press the default to encourage this to be entered. We need to report on items with missing date of acceptance as these are not captured in compliance monitoring (Manya to provide templates)
- V5.5 has additional help text to warn users against changes that could affect compliance
- Evidence of acceptance date – few sites using this as not intuitive and only wanted by minority of Schools. Document disappears if publication state is edited (comes back if you edit back to acceptance)
- We would prefer Document version to be mandatory and Public access to file to have a 'no value' option as this is often changed by library staff anyway. Manya to suggest to developers
- From V5.6 Pure plans to ingest metadata tags on publisher pdfs (Elsevier in the first instance, and other publishers as they implement). Agreed with User group request that fields can still be editable
- Several people commented that Sherpa/Romeo info (as it appears in context of deposit) is confusing for users – the colours are confused with green/gold and the statement including mention of publisher pdfs leads to people uploading the wrong version. We want to look at ways of removing the info – potentially switch off cron job and clear existing fields? Even better would be an improvement to the Romeo data that gives clearer options according to version and in context of repository deposit
- Pure is not currently looking at embargo dates in compliance as it can't be judged until REF panel is known for the output. So better to wait until items are proposed for REF and UoA is allocated. Best to focus on deposit compliance in current reporting

- **Conference proceedings not in scope of REF of policy can be marked by selection of a Technical 'exception' (**no other way to remove from scope based on ISSN/ISBN)
- Common request is a report for Directors of Research on non-compliant outputs (a 'name and shame' report). Manya showed a **report using 'Collections' reporting which includes Author names and Managing Organisation**. She will share on REF wiki
- Acknowledged that some reporting is still 'fragile' e.g. changing a field selection can mean some results are lost
- REF compliance flagging still has a known bug where dates are being derived from DOI where the DOI is Closed.

Open Access Technical Workshop
E2EOA and LOCH Pathfinder Projects, Glasgow 04 April 2016

DSpace breakout session

DSpace and the REF

Questions for JISC:
- What reporting will be available in the DSpace REF plugin?
- Will compliance be displayed in a report as well as during deposit?
- Can we see what deposits have a deadline approaching that are REFable?
- Are changes to dates recorded or just updated?
- How should we manage closed deposits? Should we just embargo the files or make private?
- How can we update existing deposits with ORCIDs?
- How will we make the final REF submission from DSpace?
- Who else apart from RGU, Abertay and SRUC are using the DSpace REF Plugin?
- RIOXX do we need to implement this DSpace Plugin as well?

Questions for SDLC implementations:
- Should Project ID and Name be included?
- Should Funder Name be added?
- Can we evaluate the DSpace-CRIS?
- Can DSpace be linked to external reporting tools? E.g. Jasper Reports?