

Notes from the OAI6 Breakout group 2: OA Repositories and Research Management Systems

Presentation of the participants:

William J Nixon, University of Glasgow:

Just got a fund from JISC to look into the interoperability between Eprints and their research management system.

Paul Ayris, UCL (University College London): Mandate for OA, library responsibility. Got commercial software to capture the metadata, prepopulate repository, as a starting point for the researchers. It will interface Eprints. Move metadata backwards and forwards. Interfacing two other systems: HR and Finance + into the annual reports. If they are not in the repository they cannot be taken into account.

Peter van der Togt, Wageningen UR. Like all Dutch universities Wageningen UR is using METIS. At this moment there is an investigation going on whether or not Metis is going to be upgraded or replaced by another CRIS. The current repository is home build. The repository contains both metadata only, metadata with link to external website and metadata with open access records. All the output is entered into the Metis system. After validation of the records by the library they are imported into the repository. On top of the repository different services are built. Most important services are Publication lists for researchers, research groups and research programmes. A new development is bibliometric analyses based on ISI citations and baselines of the Essential Science Indicators (both Thomson). With this tool they can analyse researchers, research groups and research programmes.

Saskia Franken, Utrecht University Library. Also using METIS. Using Dspace as repository. Interface between Metis and Dspace. But they are two separate systems. Attachment of full text is transferred to the repository.

Rita Vogt, Helsinki University of Technology. Using their library catalogue which has been customised to meet the requirements of RMS/repository?. But it is not SQL compliant. So everything has to be coded manually. Problem is not only technical. But that different people maintain the systems. Having different interests. Four databases projects, expertise, repository. Quality is in the repository. They import everything from the repository to the RMS. Not so many items in the repository, as the CRIS that is more populated- as it is mandatory. However, the quality is not as good because of the distributed workflow.

Friedrich Summan, Bielefeld University. Every university in Germany is doing their own thing. At Bielefeld there are several systems. Opus, CMS for publications (RMS), trying to make synergies to connect the different systems. There was a national project to improve the Opus systems by integration with RMS. Bind together the two systems. Interfaces and workflow has been discussed. Upload your fulltext and some metadata for the repository. Afterwards when the document is activated and made available it is in the RMS.

Maarten van Bentum, University of Twente. Like the other Dutch universities Twente uses METIS, but with Eprints. Project now running for two years to change the situation especially with the workflow. Distributed data entry. Hardly any good control of the data input. We had direct connection between repositories to Metis. But after Eprints the data connection has been cut because of the quality of data. But they want to connect them again. The author should do the input themselves. Library comes in as quality control and classification by an editor in the scientific group. But that is an organisational problem that they do not want to spend time doing this. How can we stimulate the scientist to do this work. End up with a pilot.

several workflow. About 20 departments. 8-9 of them have a distributed workflow where the researchers themselves are responsible for putting in the data, afterwards the data are validated by an appointed person at the department and finally it is proofread at the library, where the bibliographic metadata quality is ensured. However the detailed metadata is also a challenge for motivating users to use the repository as a repository. So they want to make enquiries into how the upload of full texts can be made easier, e.g. using SWORD.

Christian Fuhrer, Main Library University of Zürich. Switzerland is like Germany also federalistic. Every university has their own thing. The researchers are used to fill in metadata for the annual reports. From last year they had to enter the metadata and if possible upload the full text to the repository. Exported automatically to the annual report. And metadata back to the researcher web site. The concept is enter once reuse many times. Links to publisher versions. Heavy help with copyrights. List of publishers not in sherpa romeo. The level of full text coverage is rather high, around 1/3 full texts of 12000 records. 100 % will probably never be achieved.

Mogens Sandfær, Technical Information Center of Denmark. Mogens is involved in the Knowledge Exchange Project specifically looking at metadata interoperability between OAR and CRIS the goal is recommendations and the creation of a proposed exchange format. Also involved in the repository Handshake initiative/project which was created at the International Repository Manager Workshop in Amsterdam (15-16. March 2009). This project takes a rather wide view on the handshake - as transport data/objects between systems. This project will be wrapped up at the end of the year.

Andreas K. Andreou, Cyprus University Library. Running a greenstone repository, full text, videos. Not a CRIS system at the moment. But looking into getting a commercial system.

Jonas Gilbert, Gothenburg University Library One system for full texts and one bibliographic system for "RMS" imports data from PubMed. Process of trying to figure out where to go. Bibliographic system doesn't qualify as a CRIS system. The concept has not been used at the university admin. Make some integration between systems like integrated search. Simple metadata import. Sweden is discussing if they are going to use local databases for research assessment at the moment or use a central model based in e.g. ISI data.

Lee Dirks, Microsoft Research. Develop tools to support the scholarly communication life cycle and the researchers. Develop tools that interoperate. Virtual research environment. Open to repositories, Research output platform. Currently it doesn't interoperate with CRIS - but it should.

Lieven Droogmans, @mire. Comment on projects. Integration Dspace with PURE at the Flemish Research Portal. And in Dspace integrating the citations data.

Topics that were coming up several times in the initial presentations (put into three main headings):

Exchange info

- formats
- vocabularies

Integrate Systems

- Repo + RMS functions
- RMS to add repository functions to it.

Organisation and cultural issues

- collaboration to create synergy
- the stick aspect of motivation.
- political aspects - different ideas in different countries. In some countries it could...

Below is a humble effort to capture the discussion - please feel free to add.

Different views on the repository and the RMS. The RMS is at many institutions the system of the administration. The researchers do not see it as their own system. While the repository is a service for the researchers. Mixing them up might confuse or might transfer possible negative feelings about a top management RMS to the IR: But is it necessary that this happens?

Is it a good idea to merge them? In the above perspective it might be best to keep them separate. Because they have two different goals.

However, it should be easy for the researchers. You want only to enter the data one time.

Doing this is just a technical challenge. Lieven Droogmans did a project for University of Leuven where the administration wanted the publication into repository and all the other stuff into the RMS. But kept in separate systems but linked together via web services. For the researchers point of view they don't care about the repository or the RMS - they just get at web form where they can enter their data. They never see the actual system interfaces of i.e. Dspace.

It also differs from discipline to discipline. When you talk the national/regional level. Implement enough carrots to get motivated. But it took time. Classical arguments for CRIS. The usage of the repositories is very discipline specific.

Metadata is only in the repository and linked to the CRIS - but not transferring data between the systems. Instead linking metadata repository data to RMS.

It started out with Imported all metadata from a legacy system. To get the critical mass.

The CERIF format is very detailed. Cross walk difficult. This means that you will need to customize. Library runs the system, supporting. Funded by the research foundation?

Submitted and Accepted articles might be interesting for the researcher themselves, especially the younger ones. Also Theses that are a collection of articles, that might not all be published at the time of submission.

Where should the data entry start in the repository or the CRIS?

Not really clear but technically it is possible to get for instance author ID into the repositories from the beginning i.e. web service into for instance.

Is the problem that it doesn't really make sense to talk about repositories and RMS?

Well, the researchers don't care what system it is. But it is important to have data - the chicken / egg problem. So they filled in data from PubMed and other existing data that ensures content from an early state is important.

I don't want to turn my repository into a RMS, and I am quite sure that the Admin do not want to turn their RMS into a rep.

Find new ways to interlink data. How can we manage data into this?

As long as the data comes from the repository and are pushed into the own space. Then I don't care if the actually repository is interacted.

Experiences with SWORD

Microsoft integrated SWORD into Word.

User feedback from a "@mire"? study showed that researchers aren't keen on direct deposit in repository content in Word is often not mature to actual be published. It is not that dynamic.

Maybe better to first put it in to something like Virtual Research Environment. And then mature it for publication into the repository. Also introduce different forms of access - for instance own, peers, department and public.