From Research Management System to Digital Repository: Managing and storing research outputs at the University of Sydney.

Gary Browne[1]; Venkatakrihnan Balasubramanian[2] Appia; Sten Christensen[3]

This poster will demonstrate a system for transferring data and objects from the University of Sydney Research Office research information management system (RIMS) to a Library supported digital repository (DSpace). The system is currently being used for the archiving of HERDC materials. It represents a successful collaboration between the above parties and offers independence for both to use their own systems as well as recognising that the systems utilised by each have their own strengths and weakness that the other can complement.

The University of Sydney Research Office recognised that the digital archiving of research outputs was vital for access to this material, especially for reporting purposes as well as providing a historical record of the research undertaken by the staff of the University. Both partners realised that a "*one size fits all*" approach for the collection, management and reporting of this material was difficult and extremely hard to achieve. Rather than scope the perfect system it was decided to let each system excel at what it does thus achieving the best of both worlds. A system was developed where by a Library repository system (DSpace) could interface with the Research Office RIMS. We took the longer term view that the development of such a system could be further used to interface with other systems and could be coupled with our Open Access repository.

The System

The RIMS handles the initial submission, metadata and reporting of the research outputs. From a digital repository perspective much of the information required for reporting is not necessarily important for archiving purposes.

Once the research output has been entered, verified and the quality of the object checked for readability, the RIMS generates a SIP (Submission Information Package) that is sent to a shared area. The SIP comprises a folder containing:

[1] Development Programmer, University of Sydney Library IT Services

[2] Programmer, Sydney eScholarship – University of Sydney Library

[3] Sydney eScholarship Repository Coordinator – University of Sydney Library

- A Dublin Core xml file (metadata created from a mapping of the RIMS database to qualified Dublin Core)
- The objects (files)
- A content file (an inventory of the files to be loaded into the repository.)

The SIP is sent to a shared area where the Library system takes it and deposits it into the repository. The system comprises a script which predominantly uses UNIX commands and the DSpace BatchImport facility to ingest the item into a DSpace repository.

On successful ingest to the repository, a handle (persistent identifier) is generated for the item. To allow look up and reporting through the RIMS, part of the reference to the full bit stream is sent to a common file (import.map) located in the shared area. The RIMS completes the reference to construct a url to the individual objects within the submitted item[4].

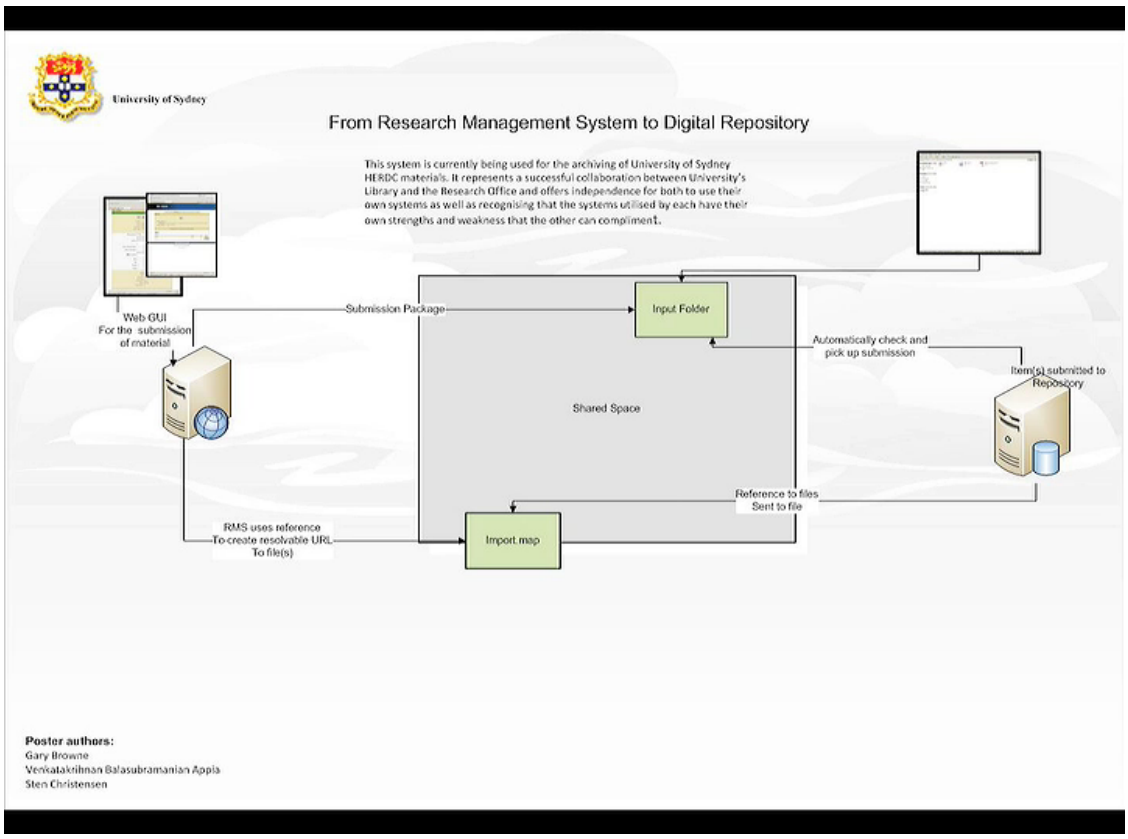

Fig 1. ***To see the process in action please place your cursor on the image and double click your mouse..***

[4] We recognize that a straight handle to the items would allow persistence, however the system was originally designed to accommodate the Research Quality Framework specifications for access to the required bitstream.

The system went “live” in January 2008 and at present we have successfully archived over 4500 items representing the University's 2001-2004 HERDC material and are, at the time of writing, preparing to do the same for the 2007 HERDC material. This material represents text, image, audio, video, software and anything that may be required for consideration for HERDC. It is also envisaged that it will be used to manage material for the upcoming ERA[5] initiative.

The Future

We see this as “Phase 1” and will enhance the system once the ERA specifications are finalised. We will also look to synchronize the system with our open access repository; the Sydney eScholarship Repository, to manage material permitted to be accessible via this medium. As well as this, we may explore the use of METS as a way of moving objects from the RIMS to the Repository[6].

Our approach has been to recognise that researchers will initially want to manage their data in their content management system of choice. Our only concern is that they be able to generate a SIP for our repository to successfully archive and disseminate their material for the long term. To this end we will use this system to manage the submission process for digital theses as well as other collections at the University of Sydney.

[5] *Excellence in Research for Australia* initiative [http://minister.industry.gov.au/SenatortheHonKimCarr/Pages/NEWERAFORRESEARCHQUALITY.aspx]

[6] The Australian METS Profiles http://pilot.apsr.edu.au/wiki/images/1/1f/AustralianMETSProfile_0.2.pdf was not fully developed during the time system was developed.