Researcher Centred Data Repository Workflows

Potential Workflows

1. DMP-derived project description, then project-end deposit of results
   - Researchers deposit a project description, possibly in response to a publisher policy on data availability. Data underlying one or more publications can be linked to them when available either in advance of or after publication. Project container can be added at a later date if the information becomes available.

2. Ad hoc deposits without project context
   - Researchers can deposit an individual data collection without a project description. Datasets without a project description are linked back to publications that they support. The data may also be deposited before or after a related publication is added to the publications repository (ROAR).

3. Publication deposit, with data as support
   - Researchers deposit a publication with data that has been archived in ROAR. The link is displayed on the summary page, and the link to the external repository is recorded in the DMP. The data is added to the project when offered by the researcher.

4. Publication deposit, with data deposited elsewhere
   - Researchers deposit a publication, which details the availability of data at an external data archive service. Using data sets as a registry, the latter is recorded in a blank data collection linking to the external data.

Abstract

Supporting the management of research data in universities requires a range of services and components, including a long-term vehicle for archiving and sharing data. Such data repositories can serve multiple functions, and ensuring they engage with researchers through appropriate workflows is a key challenge. Datauel, a new data repository at the University of East London, seeks to provide workflows for four use cases.

Technical

Both ROAR and data.uel are supported by the University of London Computer Centre (ULCC). The system was created and designed to look and feel, while ULCC built the back-end repository.

This is a many to many relationship, so the depositor can link to multiple publications in ROAR. The link is only stored in datauel, so ROAR has to look up datauel to get the links, which means it is possible to have a metadata-only record (registy) to link to an ePrint. This means it is possible to have a metadata-only record (registy) which links to the actual data.

The link is displayed on the summary page, with the colour of the destination repository designating where it is going to. This information is also in the machine readable metadata as:

`<meta name="eprints.external_link" content="http://data.uel.ac.uk:8080/api/registy" />

3. Data to external datasets (registy)
   - Links to external resources are handled by the existing ePrints link resolver, with minor customisation.

   It allows for externally stored datasets (UK Data Archive, Figshare, etc) to link to an ePrint. This means it is possible to have a metadata-only record (registy) which links to the actual data.

   The depositor inputs the URL, the title and a description of what the object is. These are then displayed in the same way as the repository links above, but with a different colour to help differentiate them.

Authors:

Stephen Grace  
Research Services Librarian  
s.grace@uel.ac.uk

David McElroy  
RDM Officer  
d.mcelroy@uel.ac.uk

Research Data Services  
University of East London

Acknowledgements:

We would like to thank Rory McNicholl and Timothy Miles-Board from the University of London Computer Centre for the excellent work developing and supporting our repositories.