DaMaRO (Data Management Roll-out at Oxford): DataBank & DataFinder

Thursday 25th October, 2012

James A J Wilson
James.wilson@oucs.ox.ac.uk
DaMaRO is building a unified RDM infrastructure for Oxford, with an institutional data archive called ‘DataBank’, and a single registry of all data produced by Oxford researchers, entitled ‘DataFinder’. The new infrastructure is underpinned by training and documentation, and an institutional RDM policy.

Project website: http://damaro.oucs.ox.ac.uk/

Project Team: Asif Akram, Bhavana Ananda, Abigail Cooke, Alexander Huber, Neil Jefferies, Prof. Paul Jeffreys, Dr. Meriel Patrick, Sally Rumsey, Dr. James A. J. Wilson
• Oxford’s in-development data archive
• Intended for datasets with nowhere else to go
• Links to ORA – the Oxford Research Archive (for research papers)
DataFinder

- Catalogue/Registry of research data
- Intended to hold records for ALL research data produced by Oxford
- OAI-PMH harvesting of data stores
- SWORD2 compliant
- DOIs assigned to records
- Interfaces directly with DataBank & ORDS
- Users can register non-electronic data
- Hierarchical structure
Possible statuses of DataFinder records

*Draft* – user still working on the record

*Submitted* – user has submitted a record for review

*Approved* – an admin has reviewed a submitted record and declared it suitable for the catalogue without further modifications

*Escalated* – an admin has reviewed a submitted record but decided it needs to be checked by someone senior who may need to call upon specialist help, due to issues regarding commercial or legal agreements, ethics, etc. A note of the issue at stake needs to be added when an admin sets a record to this status.

*Referred* – the admin has decided that the more/better information is needed before they can approve the record and therefore have referred it back to the submitter. A note of the problem and how to rectify it needs to be appended when this status is set.

*Rejected* – the admin reviewing the record has decided that there is something fundamentally wrong with it – presumably due to misunderstandings regarding the nature of DataFinder. Again, the admin must append a note, this time explaining to the submitter why the record has been rejected.
DataFinder search and browse

- Keyword & Boolean searching
- Filtering by embargo status, subject, source, etc.
- Faceted browsing
- Full record view
  - Harvested data from sources kept separate from ‘overlaid’ additions added via DataFinder
DataFinder Metadata Records

- Based on DataCite schema
  - Minimum set required
- CERIF compliant
- Additional metadata fields to meet current RC requirements
- Subject-specific metadata can be added as .xml stream
- Metadata automatically picked up from earlier tools/services
DataFinder DataReporter

- Generates standard reports
  - Institutional and departmental reports
  - Click-throughs & downloads
  - Personal data publication reports
  - Records lacking key metadata
  - Statistics for REF
- Customised reports in response to SQL queries
- Admin-only in first instance
Successes

• Institutional data management policy
• Integrating training materials developed elsewhere
• Collaboration between IT Services, Bodleian Libraries, Research Services
• Provided vision of where the University needs to be
• Pushing RDM up the University agenda
Damaro Challenges

- Getting rich, consistent metadata
- Training engagement with researchers
  - Customization of training
  - Embedding RDM training at Divisional & Departmental levels
- High-level coordination of RDM
- Developing procedures for postgraduate researchers to deposit data & metadata
- Sustainability – how do we pay for the DataBank / DataFinder services?
- Measuring impact
  - ROI only apparent over many years, and hard to isolate
  - Evidence of data reuse (particularly of data stored externally)
http://damaro.oucs.ox.ac.uk/
damaro@oucs.ox.ac.uk
Please email us to receive progress updates