

DISC-UK DataShare: building capacity for institutional data repositories

Robin Rice, Ann Green, Sally Rumsey

Abstract. This poster presentation will chart the overlapping goals, relationships, and roles of the participating data information specialists, institutional repository developers, and social science researchers at each of the four partnering institutions of the DISC-UK DataShare project in relation to the data life cycle. The presentation also will introduce a new model of data sharing and archiving for UK research institutions, emphasising collaboration and sustainability throughout the data life cycle.

Keywords: data sharing, institutional repositories, data lifecycle, collaboration, stewardship.

The DataShare project's overall aim (<http://www.disc-uk.org/datashare>) is to contribute to new models, workflows and tools for academic data sharing within a complex and dynamic information environment which includes increased emphasis on stewardship of institutional knowledge assets of all types; new technologies for doing e-Research; new research council policies and mandates; and the growth of the Open Access / Open Data movement.

The project seeks to build capacity of institutional repositories in UKHE to respond to the unique requirements of research datasets. The project is based upon the existing informal collaboration of DISC-UK members (Data Information Specialists Committee-United Kingdom) for improving their data libraries and models of social science data support. The DataShare project will bring academic data libraries in closer contact with institutional repository managers and will develop new forms of cooperation between these distinct groups of information professionals within academic environments. It will identify training needs for both groups, particularly librarians unfamiliar with support of research data.

Based on a distributed model, the DataShare participants at four institutions (Edinburgh, London School of Economics, Oxford, and Southampton) are each responsible for the work on their own repositories, yet experience, support and knowledge are shared in order to increase levels of success. This network of institution-based data repositories develops a niche model for deposit of 'orphaned datasets' currently filled neither by centralised subject-domain data archives/centres/grids nor by e-print based institutional repositories (IRs).

The advantage for the broader community is to provide exemplars for a range of approaches and policies in which to embed the deposit and stewardship of datasets in institutional repositories. These will be demonstrated using the three main open source repository solutions: EPrints, DSpace and Fedora. Policy and scoping documents, depositor agreement forms, preservation planning, software code and templates for institutional data audits will all be shared with the community.