

Opportunities for ‘data intensive’ social research are growing but funding for data management remains a challenge.

 blogs.lse.ac.uk/impactofsocialsciences/2014/03/25/research-data-management-strategy-funding-whyte/

3/25/2014

The Digital Curation Centre have been vigorously involved with identifying core components of an effective institutional policy to improve research data management. [Angus Whyte](#) runs through the opportunities and challenges in business planning for RDM. Funding and sustaining such services requires a clear vision of what better data management will do for the institution, its researchers, and the broader community. The key is to avoid unintended outcomes of poor design of services or inadequate resourcing.



At some point in the last five years “research data” acquired capital letters and a “management” appendage to become Research Data Management, inevitably abbreviated to RDM. This has emerged as funding bodies coordinate policy demands on universities to help researchers do this effectively. It also reflects an international government push towards open data, as illustrated by the G8 Science Ministers [statement](#) last year.

High-level concerns for RDM also come from the research community as it tries to address a ‘growing crisis of reproducibility’, as the Royal Society’s [Science as an Open Enterprise](#) 2012 report put it. High-impact journals including *Nature* now require published articles to have supporting data available. These are not just STEM disciplinary concerns; e.g. one of the main strands of the ESRC’s [National Data Strategy](#) is to implement the Royal Society report’s recommendations on sharing personal information for research.

The strategy also sees growing recognition of social factors across disciplines, and consequently more interdisciplinary research. Just as that pushes against domain-based cultural barriers to data sharing, opportunities for ‘data intensive’ social research are growing. These can be seen in the opening up of administrative data, and the growth of ‘research infrastructures’, such as [CESSDA](#), [European Social Survey](#), and [SHARE](#), which expand the scope and richness of reusable data. But much of this isn’t that new. Institutions are already making headway to tap into these opportunities and nurture their local RDM infrastructure. What changes are institutions making now? With what gains for them, their researchers, and for wider society?

I’ll declare an interest in these questions; in the Jisc-funded [Digital Curation Centre](#) (DCC) we help institutions build RDM support through our [tailored support](#) programme. We find that librarians, research office and computing managers have no trouble recognising the basic rationale for developing services; help to exploit data, and to comply with policies requiring public data sharing. Nobody wants to burden researchers with time-wasting tasks. The challenge is to avoid that happening as the unintended outcome of poor design of policies or services, or inadequate resourcing. And resourcing is needed if data management is to gain recognition in its own right, both as a research activity and a career option for support staff.

There is already evidence that organisational support encourages researchers to share data by depositing it in a public repository. For example a study of [determinants of data sharing](#) published last year found that skills development and organisational support for RDM was the single most important factor.

So what institutional RDM services are taking shape? Many begin with RDM policy (see our [list](#)). These are statements of principle and good practice, often making a clear commitment to develop a support service. From our experience with more than 40 institutions through our programme and Jisc funded projects, distilled in a [How-to guide](#), we identified seven ‘core components’ of RDM services to implement policy; business planning, data management planning, active data management, data selection and handover, data repositories, data catalogues, and the guidance and training to build skills in each of these. Many institutions start building such capabilities by

providing [online guidance](#).

A 2012 [survey](#) of 81 UK university libraries found top RDM service priorities to be policy development, establishing advisory services, data repositories, guidance on reusable data sources and citation, and training. Loughborough University's more recent 2013 [survey](#) found that, among 38 responding institutions, 16 had an approved RDM policy, with another 15 in draft. Services to implement policies were being developed by 25, with only 6 having live services.

DCC is currently surveying UK institutions to get a fuller picture. We aim to identify how far they have got in each of the areas already mentioned, where they have invested in staff, what they regard as the main barriers, and where they see the opportunities for cross-sector collaboration. Findings will be reported in mid-May. What of the future? I expect we shall see different institutional models emerging. Some will be centralised around an institutional data repository. Others will likely be networks of e-research specialists, sharing some pooled resources including storage and an institutional data catalogue but otherwise doing their own thing.

Our language of 'core service components' might suggest that all institutions need do is assemble these into a gleaming machine, then lower it into the engine room of the great research support ship, launch it and watch their researchers pile on board, and reap the benefits. Nobody expects that uniformity or instant impact, although there are substantial [returns on investment](#) in disciplinary data repositories like the UK Data Archive (estimated at 2.5 to 10 fold). Nevertheless the economics of institutional RDM services are not so rigorously established that financing them is easy. It is straightforward to identify costs and benefits of RDM. Standards for *quantifying* them are more difficult and the benefits are not so well established for research fields that do not already routinely share data.

Business planning for RDM is therefore a challenge. Funding and sustaining such services requires a clear vision of what better data management will do for the institution, its researchers, and the broader community. For the institution, RDM benefits include improvements to the research environment, countable for (the next) REF purposes, and seen by some institutions as a source of competitive advantage. For researchers, efficiency savings are countable (e.g. from centralised backup), as are the benefits of being able to pool datasets for retrospective analysis, and the knock-on effects of data sharing on [publication productivity](#). Direct advantages to those beyond academia are also identifiable. Studies with University of Bath for example identified a range of [benefits to external partners](#) and communities, e.g. confidence in the handling of personal or commercially sensitive data among research participants and partners, and better access to reference datasets.

What about researcher take-up? Support for Data Management Plans is a clear and current need. More broadly, I expect demand for advice to be highest where researchers take on collaborative projects that stretch the normal boundaries of their fields. And as projects with DMPs get funded, if even a fraction keep to their plans we can expect strong demand for help to find tools and repository services that will really make their data work for them, as well as to mine what's already out there.

Note: This article gives the views of the author, and not the position of the Impact of Social Science blog, nor of the London School of Economics. Please review our [Comments Policy](#) if you have any concerns on posting a comment below.

About the Author

Dr Angus Whyte is based in the Digital Curation Centre, funded by Jisc to help UK Higher Education institutions develop support for research data management. He had 10 years post-doc experience as a social informatics researcher before that.

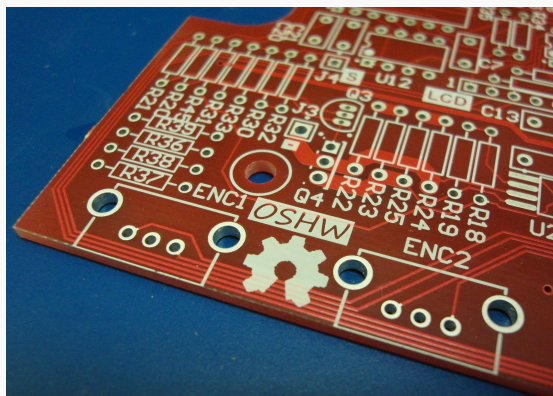


Image credit: Altzone (CC BY)

- Copyright © The Author (or The Authors) - Unless otherwise stated, this work is licensed under a Creative Commons Attribution Unported 3.0 License.