Molloy, L., Hodson, S., Poschen, M., and Tedds, J. (2012) *Gathering evidence of benefits: a structured approach from the JISC Managing Research Data Programme.* In: International Digital Curation Conference, 14 – 16 January 2013, Amsterdam, Netherlands.

Copyright © 2012 The Authors.

http://eprints.gla.ac.uk/86326/

Deposited on: 09 October 2013

8th International Digital Curation Conference
January 2013

Gathering evidence of benefits: a structured approach from the Jisc Managing Research Data programme

Laura Molloy,
Researcher, Jisc MRD Evidence Gatherer,
Humanities Advanced Technology and Information Institute (HATII),
University of Glasgow

Simon Hodson,
Programme Manager, Managing Research Data Programme,
Jisc

Meik Poschen,
Research Associate, Jisc MRD Evidence Gatherer, MiSS Project Researcher,
Manchester eResearch Centre (MeRC),
University of Manchester

Jonathan Tedds,
Senior Research Fellow (D2K Data to Knowledge Group, Dept of Health Sciences),
Jisc MRD Evidence Gatherer,
University of Leicester

Abstract

The work of the Jisc Managing Research Data programme is – along with the rest of the UK higher education sector – taking place in an environment of increasing pressure on research funding. In order to justify the investment made by Jisc in this activity – and to help make the case more widely for the value of investing time and money in research data management – projects and the programme as a whole must be able to clearly express the resultant benefits to the host institutions and to the broader sector. This paper describes a structured approach to the measurement and description of benefits provided by the work of these projects for the benefit of funders, institutions and researchers. We outline the context of the programme and its work; discuss the drivers and challenges of gathering evidence of benefits; specify benefits as distinct from aims and outputs; present emerging findings and the types of metrics and other evidence which projects have provided; explain the value of gathering evidence in a structured way to demonstrate benefits generated by work in this field; and share lessons learned from progress to date.

Introduction

The work of the Jisc Managing Research Data programme aims to produce an improvement in the way research data is managed throughout the research lifecycle in a range of institutions across the UK. This work is – along with the rest of the UK higher education sector – taking place in an environment of increasing pressure on research funding; in order to justify the investment made by Jisc in this activity, projects and the programme as a whole must be able to express clearly the resultant benefits to the host institutions and to the broader sector. This not only justifies funder investment but is essential to make the case for sustainable investment within institutions in a period of contraction of resources (Whyte and Tedds, 2011).

The second Jisc Managing Research Data ('MRD') programme, running from October 2011 to July 2013, has at its core a set of seventeen relatively large projects, each lasting between eighteen and twenty-one months, focussing on the development of research data management support services and infrastructure in higher education institutions ('HEIs'). For the most part, and notwithstanding institutional specificities, the projects are developing and implementing relatively comparable components of research data management services, comprising: policies and roadmaps; guidance materials and training activities; systems for managing and storing active data; guidance and processes for appraisal and selections; processes and systems for deposit and exchange of metadata; institutional data repository platforms and/or data catalogues; and a business case to sustain the research data management ('RDM') support service.

The benefits that accrue from this work can be analysed in terms of the recipient (institution, researcher/research group, support service, broader research community) and in terms of timescale. Some benefits can clearly be quantified in terms of costs avoided or improved efficiency, although gathering the evidence and determining value – especially within the timescale of a relatively short project – may be challenging. Other benefits seem less susceptible to quantification and may depend on the project scope (pilot groups / institution-wide) or stakeholder, but may be compellingly presented in qualitative narratives. There is also a distinction to be made between the tangible benefits a specific project may bring to an institution, research group or discipline, and the broader benefits which it is argued may emanate from improved availability of research data.

The second Jisc Managing Research Data programme has undertaken a structured programme of activity to develop, alongside funded projects, an appropriate model for identifying and describing the benefits of the individual projects and thereby to synthesise and articulate the benefits delivered by the work of the programme as a whole. It has been conducted in such a way (including wide use of blogging) that projects are aware of related work and evaluation of benefits and evidence in other projects and may therefore adapt and benefit from these during the lifetime of their own project. The relative alignment between the objectives and activities of the projects promises a similar alignment between benefits and the type of evidence to be gathered in substantiation, and therefore a healthy prospect for synthesis across the programme. We will be able to offer more concrete conclusions once the programme activity is completed and final benefits statements have been collated and analysed.

The MRD Approach to Benefits and Evidence

Evidence of the benefits of individual projects, and the overall Jisc Managing Research Data programme, has been tracked since the programme was launched in 2009. The eight research data infrastructure projects in the 2009-11 ('MRD01') programme each produced a benefits case study and these were synthesised into a report (Beagrie, 2011). This was a useful exercise to help begin to build an evidence base for the programme, as well as a constructive way to improve the tools employed in scrutiny of the evidence identified. However, it was not always possible to capture valuable learning experiences during the project lifetimes and it was harder to synthesise benefits across the programme given that – in this more exploratory and experimental phase of work – approaches varied more considerably between the projects than in the second programme ('MRD02'). The identification of appropriate evidence to be gathered was also found to be challenging for projects that were operating in relatively uncharted waters and whose outputs, in many cases, were prototype processes and software. Nevertheless, significant progress towards a common understanding of what the development of RDM services required was achieved; this allowed the second programme to develop a more coordinated and effective approach.

Projects funded in MRD02 were more broadly aligned in their activities and the approach to providing evidence of benefits has been more structured, coordinated and deliberate. To achieve this, MRD02 funded three part-time 'Evidence Gatherer' ('EG') posts at 0.2 FTE each, specifically to work with projects towards articulating a narrative of the demonstrable benefits resulting from each project and to correspond across projects in the programme where appropriate. Each EG is a researcher with significant experience in this area as a member of a project team on MRD01.

A Tailored Approach

At the first programme-wide event for the second phase of the programme, the projects were introduced to the Keeping Research Data Safe ('KRDS') Benefits Analysis Toolkit, developed during MRD01.[1] Projects then informally published a statement indicating the benefits they anticipated emerging from their work. Whilst maintaining an overview of the programme as a whole, each EG has worked particularly with five or six projects. This feature has been an important element in the delivery of the evidence-gathering work, encouraging communication between named contacts and aiming to provide a continuous source of reference and assistance for the project. Each project, as part of the requirements for funding, developed a statement comprising a list of benefits agreed between the project and their EG. For each specified benefit, the project then identified appropriate and specific evidence which could realistically be provided. Cognisant of benefit mapping approaches,[2] the emphasis here was on keeping the process lightweight and achievable within the relatively short timescales of the projects and within the limits of the resourcing of each project.

---

[1] Keeping Research Data Safe Benefits Analysis toolkit: http://beagrie.com/krds-i2s2.php

[2] E.g. as in MSP and in HM Treasury. (2011) The Green Book: Appraisal and Evaluation in Central Government.

Effort was required in the early stages of this work to draw on the benefit mapping approaches mentioned above and then to set out definitions of 'output', 'benefit' and 'evidence' in the context of the programme. For the purposes of the evidence gathering function, the *output* was defined as 'something that the project is going to make, produce, put in place or that it otherwise aims to deliver', as specified in the project plan; the *benefit* as something that can be identified by asking, 'What does this help us (the institution / researchers) to do better?'; and *evidence* as 'specific, clear metrics (quantitative measures) and specific, clear qualitative evidence such as narratives and short case studies, all of which support or prove the benefit' (Molloy, 2012). This information is to be presented in a specific report delivered alongside or included in the usual Final Report of the project, but projects were also encouraged to blog and tweet about this work as it happens in order to help the flow of information across the programme.

Clear Direction, But Not Too Much

The EG team and programme manager encouraged a focus on gathering reliable evidence for a limited number of benefits, with the expectation that these would be relatively consistent across the programme. Distinct benefits emerging from a particular project alone were also included where appropriate. In this first attempt at modelling a programme-specific approach, however, it was important to allow themes to emerge from project thinking, rather than being prescribed by the programme team at the outset. This approach allowed the projects to exercise considerable influence over the themes arising and resulted in community-generated benefits which are deeply supported and extensively evidenced by the programme's work in multiple ways, whilst still limiting the number examined *per project*, to retain clarity of focus. This approach did, however, involve more complexity than prescribing a limited number of benefits from which projects could choose.

Projects were asked to set mechanisms in place at an early stage in the project to track quantitative metrics (e.g. website visits, usage statistics of web resources, number of enquiries for support received, number of data management plans created and approved, attendance at RDM training events organised, etc.) in order to provide an early benchmark for comparison at project end. Adaptations of the Digital Curation Centre's (DCC) CARDIO tool[3], or bespoke questionnaires, will be used by some projects to provide benchmarks against which progress in such areas can be measured. Projects were also, however, asked to enrich the detail and narrative of their benefits report with qualitative evidence such as short case studies of how the work of their project has enhanced the research practice of a particular researcher, group or institution. These might include, for example, qualitative accounts of positive impact brought about by improved data security or data sharing, or an assessment of more efficient RDM processes. A workshop was held in November 2012 to assess interim findings, to coordinate activity and progress towards a synthesis of the benefits emerging from across the programme as a whole.

---

[3] DCC's CARDIO tool is available at http://cardio.dcc.ac.uk/ but the shorter 'Pulse Check' quiz based on the full tool has also been popular with projects and can be found at http://cardio.dcc.ac.uk/quiz/. Many projects have adopted "CARDIO lite" as a nickname for the Pulse Check version.

Initial Findings

The broad alignment in activities and outputs across the projects of MRD02 coincides with a similar level of alignment in benefits and evidence emerging across the programme. Some benefits follow from specific outputs (e.g. the development of guidance materials, the production of training events, or the introduction of a particular tool) while others follow from the ensemble of outputs which comprise the RDM support service (including the development and implementation of policies, roadmaps, technical infrastructure, etc). It will be important – and instructive – for projects and the programme-level synthesis to be clear about these relationships.

Benefit Categories and Relationships

Our approach has allowed us to begin modelling an understanding of the various benefits as characterised by the work of the programme and how they relate to each other. It was hoped that as this modelling developed, improved practice in research data management – the overall aim, after all, of the work of the programme – would emerge as a central, unifying component of the model; this has, at the time of writing, proved to be the case.

At the time of writing, the work of the programme, including the evidence-gathering effort, is still underway. Final benefits statements and the resulting analysis is therefore not yet available. However, we have attempted to begin the characterisation of benefits emerging from the various projects of the programme and their interrelationships.

The model provided (see Figure 1, below) shows an attempt to demonstrate this. Each dark blue element has been cited as a benefit by the projects; the EG team has worked to sort these into

- benefits contributing to improved RDM practice;
    - elements thereof;
- benefits resulting from improved RDM practice;
- possible further benefits which may result from these.

We currently identify three main benefits which can be seen to make major contributions to improved RDM practice:

- Raised RDM awareness and understanding of researchers and institutional support staff;
- Improved RDM skills;
- Improved institutional support for RDM.

These three main benefits are comprised of a number of contributory benefits, each resulting from project interventions or outputs. Projects are asked to supply specific evidence for each of these contributory benefits.

The pale blue boxes indicate potential further benefits of the work of the projects. This is not speculation; rather, it is based on such further benefits identified as likely to result from the project work.
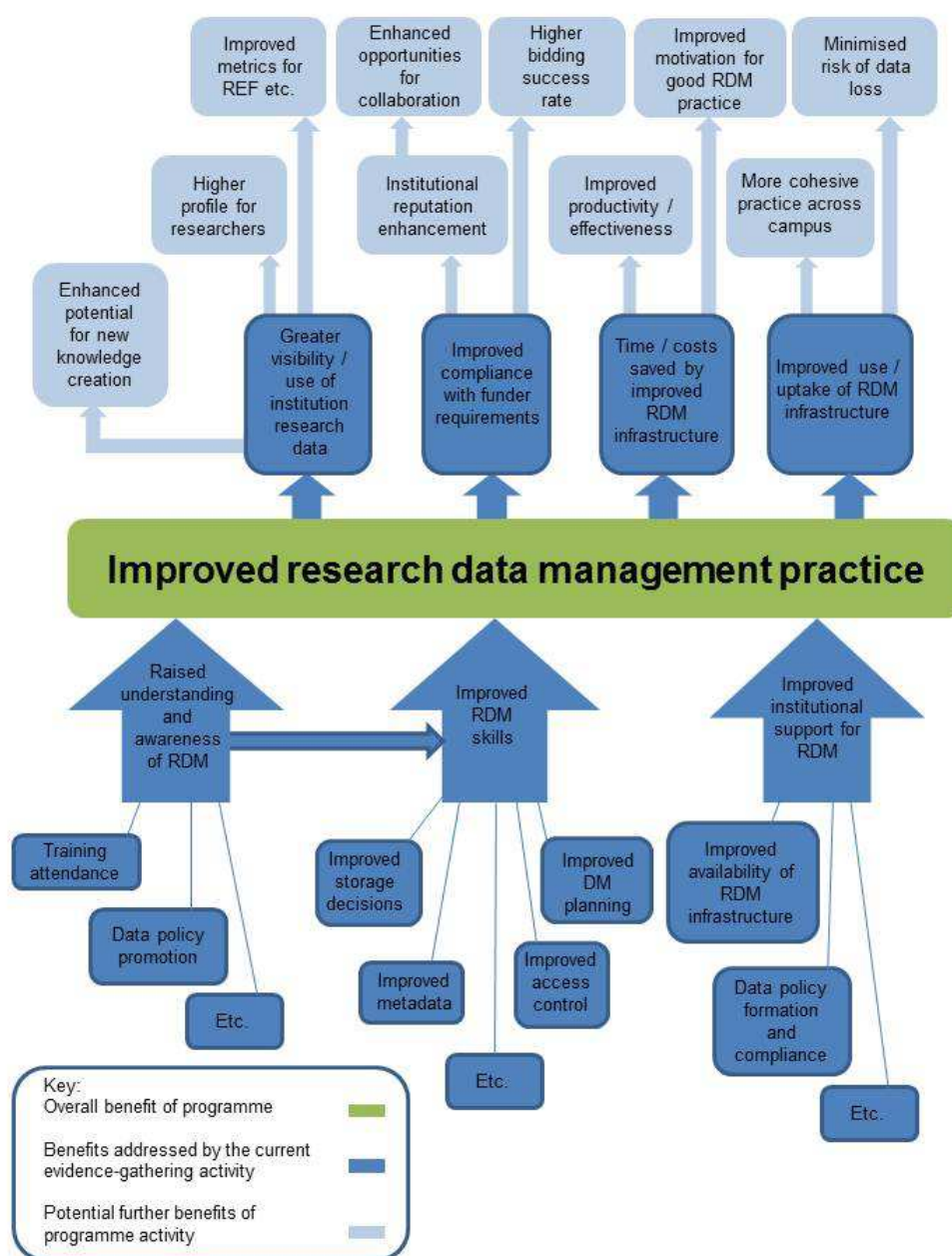
Figure 1. Interim model of types of benefit identified by the work of the programme, and their interrelationships.

Naturally, there are further relationships between these categories, and elements could recur in more than one location. The most obvious further relationship is indicated on the model: the likelihood of raised RDM awareness to contribute towards improved RDM skills. Many other such relationships are emerging from the projects' benefits work; for the sake of clarity, these are currently not indicated in the diagram here but we hope to produce more sophisticated modelling in the future, based on the benefits reports received at programme end.

Discussion

Benefits for the research sector as a result of improved RDM practice are numerous, and widely discussed elsewhere. They are represented in the current model. However, in line with the overarching aims of the MRD programme, the model also attempts to indicate the benefits of improved RDM practice for the university or institution. Whilst evidence (e.g. Molloy, Snow, 2012) suggests that many researchers strongly identify as working within a particular discipline rather than institution, it is inescapable that the institution is often the provider of the infrastructure which makes daily research work possible. The increasingly competitive nature of research funding also drives institutions to consider their competitiveness in economic as well as reputational terms. It is unsurprising, then, that in this context, projects have identified many benefits which support research practice in the institutional context. Examples include reputation management; greater visibility of the institution's research data in order to make possible greater use and reuse; improved compliance with funder requirements which may improve bidding success; and the reduction of data loss due to inadequate procedures and back-up. Although for a number of projects it presents a challenge to produce clear numbers, the cost to the university of data loss, or the risk of data loss as a potential cost can in principle be quantified, as can the benefits of increased adoption of more secure and reliable storage. The numbers of researchers with improved practice using secure, centrally-provided storage or the equivalent can be used as a performance indicator for RDM services; here, mitigating the risk of data loss is a key benefit.

The enhancement of links with industry or other collaborators may be another of these further benefits. In the case of one of the MRD projects, the university has identified the RDM infrastructure being developed as something which will facilitate collaborative projects with industrial partners. Case studies will furnish qualitative evidence of this benefit. In the model, enhanced potential for collaboration is shown to flow from the enhanced institutional reputation, which in turn flows from enhanced compliance with funder requirements. It could be argued, however, that use of an effective and fit-for-purpose RDM infrastructure also contributes to the enhancement of the institutional reputation. This is an example of the type of intricacies emerging from this work which we hope to address in the programme-end analysis.

Projects also seek to provide evidence of efficiencies that may flow from the implementation of improved processes and systems. Examples include more effective processes for developing data management plans ('DMPs'), effective transfer of information between systems and easier deposit of research data to an archive or repository. In many cases it will be possible to make reasonable estimates of time-savings.

Another particularly important intended benefit to flow from improved RDM is the greater visibility and reuse of institutional research data assets. All the MRD02 projects are implementing data repositories of some stripe that will hold data and/or metadata records produced by research projects conducted in the institution.

Evidence

Projects are gathering a mixture of qualitative and quantitative pieces of evidence for each benefit they claim. Quantitative evidence of benefits is challenging to obtain at this stage in the development of institutional RDM services, not least because the projects are primarily in the process of implementing new systems, processes and activities. Projects are gathering quantitative information to provide a benchmark and making recommendations to demonstrate the benefits over time of the embedded service which will result from the work of the project. As an example, we have already mentioned that the projects are implementing or enhanced a data repository of some kind; when attempting to quantify the use of the institution's research data, the metrics most commonly to be applied are: the number of research datasets published with enhanced metadata; the number of data deposits; the number of downloads of datasets; the number of citations to datasets. Typically projects will enhance their accounts with case studies providing researcher perspectives on user experience and likely benefits from the system implemented.

The final and most common area in which projects will provide evidence of benefits relates to the uptake and impact of guidance materials and training. Many projects have developed guidance and training materials in support of an institutional research data management policy. Measuring the effectiveness of training provided is therefore also a common activity. Broadly speaking, two approaches are being taken towards two discrete but related aims. One is to provide evidence of the uptake of training using quantitative indicators: the number of training courses and events run, attendance numbers and indicators obtained through feedback. This evidence will help demonstrate improved *awareness* of RDM, whether or not training has an effect on subsequent practice. Qualitative feedback from training, however, can provide evidence of the *impact* delivered by the training by gathering information on what was learned or absorbed by trainees. This information will be particularly rich if supplemented by follow-up questionnaires to assess the impact of training on subsequent practice, although this is unlikely to take place within the relatively short timescales of the programme. However, projects keen to measure the effectiveness of interventions over time plan to implement annual researcher surveys which will provide a measure of awareness of RDM policy and practice and the increase in RDM skills obtained through the training offered. Collated at the programme level, such information is likely to be susceptible to 'shadow pricing' techniques by which it will be possible to demonstrate the significant economic value of Jisc's investment.

Challenges and Solutions

Challenges we have encountered in undertaking this exercise include a frequent conflation of, particularly, what constitutes an output and a benefit. It was quickly apparent that even with an apparently simple and lightweight approach, results were complex and intricate. As noted above, benefits are frequently linked; they may also be hierarchical. Further, a benefit may be supported by several pieces of evidence; a piece of evidence-gathering undertaken by a project may supply more than one type of evidence. None of this in itself is problematic from the point of view of those attempting to describe and promote the benefits of the work of the programme and its projects; however, these intricacies may add to the complexities of accurately presenting evidence.

Projects also vary in their experience of gathering qualitative or quantitative evidence. This said, some projects *are* able to make estimates of the cost of data loss, or of the value (in terms of investment) of data currently held on sub-optimal storage. However, there may be an understandable reluctance to make such findings publicly available. Similarly, tangible metrics of the use of DMPs or secure storage might not be available in the project lifetime.

The benefits of better practice of RDM, implying the evolution of users' practice, are often approached in the form of qualitative, longitudinal case studies and it can be argued that this in-depth view is necessary to capture the intricacies of change. At the same time, a longitudinal survey of researcher views – for example, that provided by University of Oxford's DaMaRO RDM Survey (Patrick, 2012) - can be a powerful way to show awareness, engagement and opinions pertaining to better practice and to potential or actual changes in practice. However, this is not an exercise that can be completed in the context of an eighteen-month programme of activity.

Lessons Learned

It will only be possible to appraise the success of the MRD02 Evidence Gathering activity at the end of the programme, when project reports have been submitted and the overall view of synthesised by the EG team. At this early stage, however, there are grounds to believe that the approach taken is supporting our aim of coordinating the gathering and articulation of evidence across the programme.

A crucial point in the gathering evidence of benefits approach has been to start working with projects as early as possible in the lifetimes of MRD02 programme and projects. This enabled the Programme Manager and Evidence Gatherers to establish a dialogue with the projects and provide the necessary time to think about how to define benefits and gather evidence. However, although some institutions have adopted and embedded benefits management methodologies and were familiar with elements of the approach taken here, other projects found the work more challenging. On reflection, the evidence-gathering activity might have provided even more and earlier support for projects in this activity, specifically by conducting exercises in benefits mapping approaches.

It has been stressed throughout that the evidence gathering exercise should be valuable for projects and their institutions as well as for the programme. For each project, the evidence gathered will reinforce the business cases being developing for an ongoing RDM support service. In this respect, significant lessons were learnt from MRD01. All benefits case studies produced by projects in MRD01 were closely coupled to the further requirement of delivering a business case based on a model template. The activity of benefits evaluation is a necessary step towards developing a compelling business case for an RDM support service. However, in MRD01 the projects did not always have the high-level support to make these initiatives effective at the time. This led to the new requirement for projects in MRD02 to include explicit senior management support (e.g. as the project sponsor and chair of the steering group) from the start of the project, i.e. the proposal phase onwards. This – and the greater prominence nationally of the RDM challenge – means that all current projects are preparing to develop from project to sustainable service (albeit with varying service models and different amplitude as reflects the institutional approach and

priorities). It also encourages projects to think in the long term, which for the benefits side means to include an ongoing activity of evidence gathering for RDM services going forward.

The benefits workshop conducted in November 2012 allowed projects to present and discuss their proposed benefit and evidence sets. It also provided an occasion to reflect on the MRD02 approach to evidence gathering, to consider findings and lessons learned to date. Our interim view is that the Evidence Gathering activity has substantially aided the projects in identifying the most important benefits that their work will realise. Above all, the activity has helped the projects focus on an approach to providing gathering which is realistic and tractable. A principle of the approach taken was to respect institutional and project diversity and not to impose, from the outset, a highly circumscribed set of benefits to be studied. Nevertheless, as discussed, there has been a relatively high degree of convergence. The alignment of programme activity around the development of RDM services contributed to this, but the coordinating work of the EGs has been the most significant factor. The greater emphasis placed in MRD02 upon identifying what evidence can realistically be provided by projects – and what evidence gathering should be written into ongoing activities – has allowed us to identify categories of benefits and substantiating evidence which we hope will be useful to future investments in RDM.

Conclusions

The ultimate objective of our MRD02 evidence-gathering activity is to synthesise a clear, useful and substantial narrative providing clear evidence of the benefits of the programme's work. When compiled and synthesised at the end of the MRD02 programme, this evidence can contribute towards the case for sustained investment in RDM. At the institutional level, the evidence gathered may be used to substantiate the business case for developing RDM support services. Providing tangible evidence of benefits, defining the beneficiaries and articulating how the development of RDM infrastructure and services supports a university's mission is crucial to obtaining support from senior management for any ongoing institutional commitment.

There is an emerging body of work which seeks to show the economic benefits and impact of services providing access to public sector 'reference data' and research data (see, e.g. Houghton, 2011; Houghton, Beagrie, 2012.)[4] It would currently be impossible to apply these methods to the emerging research data services available both internally and externally to universities, but a coordinated approach to benefits analysis and evidence gathering can provide the sector with advocacy tools for investment in improving RDM support and pave the way for more work which can consider broader economic impact. This will be beneficial in making the case for further sustained investment at a national level. We hope sharing these emerging findings may be instructive for any comparable or future work in this area seeking to apply a systematic approach to gathering evidence of benefits achieved.

---

[4] In addition, Jisc has commissioned two further studies assessing the economic impact of the Archaeology Data Service and the British Atmospheric Data Service.

References

[report] Charles Beagrie Ltd. (2011). Benefits from the Infrastructure Projects in the JISC Managing Research Data Programme. Retrieved from http://www.jisc.ac.uk/whatwedo/programmes/mrd/outputs/benefitsreport.aspx.

[report] Houghton, J. (2011), Costs and Benefits of Data Provision. Retrieved from http://ands.org.au/resource/cost-benefit.html.

[report] Houghton, J., Beagrie, N. (2012) Economic Impact Evaluation of the Economic and Social Data Service: http://www.esrc.ac.uk/impacts-and-findings/impact-assessment/economic-impact-evaluation.aspx.

[blog post] Molloy, L. (2012). Evidence Gathering: The Field Guide. Retrieved from http://mrdevidence.jiscinvolve.org/wp/2012/08/30/284/.

[journal article] Molloy, L., Snow, K. (2012). The Data Management Skills Support Initiative: synthesising postgraduate training in research data management. International Journal of Digital Curation, 7 (2). pp. 101-109. ISSN 1746-8256 (doi:10.2218/ijdc.v7i2.233)

[blog post] Patrick, M. (2012) DaMaRO Survey Results – Research Data Management Training for the Sciences. Retrieved from http://blogs.oucs.ox.ac.uk/damaro/2012/11/21/damaro-survey-results-research-data-management-training-for-the-sciences/

[DCC briefing paper] Whyte, A., Tedds, J. (2011). Making the Case for Research Data Management. Retrieved from http://www.dcc.ac.uk/resources/briefing-papers/making-case-rdm