# Supporting Researchers with their Research Data Management: Professional Service training requirements: report

**Dorothy Byatt,** JISC DataPool Project, University of Southampton Library, University of Southampton, UK, **De Luca,** Federico, Postgraduate Researcher, Faculty of Social and Human Sciences, **Harry Gibbs**, Academic Liaison Librarian, University of Southampton Library, University of Southampton, **Meriel Patrick** Analyst and Training Officer, DaMaRO Project, University of Oxford, **Sally Rumsey**, Digital Collections Development Manager/DaMaRO Project, Bodleian Libraries, University of Oxford, **Wendy White**, Principal Investigator, JISC DataPool Project, University of Southampton Library, University of Southampton

## Introduction

The *Science as an open enterprise* report from the Royal Society (2012) confirmed the importance of data in the conduct and communication of research. Its second recommendation highlighted the role universities and research institutes should play in supporting this through “developing a data strategy and their own capacity to curate their own knowledge resources and support the data needs of researcher” (Royal Society, 2012, p.72). The Research Councils UK (RCUK, 2012) welcomed this report and confirmed its commitment to improving access to the outputs of its research as had been demonstrated earlier in its own *Common principles on data* *policy* (RCUK, 2011). The development of data repositories, such as DRYAD (<http://datadryad.org/>) and Figshare (<http://figshare.com/>), shows that researchers have engaged with these views. Good data management is required throughout the research lifecycle to ensure that data are of good quality and re-usable and this in turn requires that there is support within the institution for the researcher. Organisations such as JISC have enabled research into the infrastructure – technological and organisational, required to support research data management and dissemination. This report looks at the work carried out as part of the DataPool project to identify the training needs of the professional staff that support the researcher in various aspect of data management.

Through the JISC funded Institutional Research Management Blueprint Project (IDMB) the University of Southampton developed its 10 year blueprint (Brown et al, 2011) for building the required infrastructure. It did this by investigating what researchers were currently doing with their data and what they thought they required. As well as the blueprint, the IDMB project also developed a draft research data management policy to underpin this work. In *DataPool: Engaging with our Research Data Management Policy* White & Brown (2013) detail how this draft policy was refined and approved. The policy on its own is insufficient but is an important step in enabling the development of the supporting infrastructure, both technological and personnel. The training strand of the DataPool project included an assessment of professional development requirements for staff supporting researchers in managing their data throughout the research life cycle. This report will focus on the investigation undertaken to assess the level of expertise in the relevant support staff groups, identify the training needs of those staff and consider what networks need to be developed to enable collaborative support of researchers in the area of research data management. It will report on the results of the survey carried out at the University of Southampton.

At the start of the project, the University of Southampton had recently completed an institutional restructuring, moving away from a three to an eight Faculty structure, and with common services provided through central professional services with some staff embedded in Faculties. Certain roles and responsibilities had been re-assigned and therefore were still bedding down in the organisation. As a result the project focussed on working with the Library, iSolutions and the Research and Innovation Services (RIS), while being aware that staff in Finance and others embedded in the Faculties also have a role to play and would need to be included in any training developed as and when it was possible. The development of the University policy on research data management had already resulted in collaborative work between these three key groups. This was further built on through the writing of the research data management website (<http://www.southampton.ac.uk/library/research/researchdata/>) with each group providing the content of the pages most relevant to their support and area of expertise (see Byatt & White, 2013). This work had involved individuals from the Professional Services at a top level and included consultation with researchers, but it had not necessarily enabled the drawing together of all the staff who worked directly with researchers on a daily basis. It was this group who were the most likely to be asked for advice or assistance. As well as developing training for the researchers (see Byatt et al, 2013), those supporting them will also need training to develop their knowledge and expertise of research data management.

## Approach

In the University of Southampton, work on research data management support began with the IDMB project. However the focus of that project was on finding out the needs of the researcher rather than those that supported them. In the DataPool project there was an equal emphasis on looking at how to develop the skills and training of both researchers and the staff that support them.

Key to investigating the knowledge and expertise held within the professional staff was the need to understand what was required by the researcher and when this was in the research life cycle. The IDMB project questionnaire results reviewed and informed our approach. Also some work in this area had already been carried out by projects such as the JISC and RIN co-funded *Research Data Management Skills Support Initiative* (DaMSSI), led by Digital Curation Centre (2011) investigating the use of the Vitae’s (2010) *Researcher Development Framework (RDF)* in this context. Therefore, part of the scoping work was done using the Information literacy lens of the Vitae Researcher Development Framework (CRAC et al, 2012) on its release along with the research lifecycle steps outlined in Brewerton (2012, p.104). This was used as mapping tool to draw together the training needs of the researcher and the support staff into one [training matrix document](http://datapool.soton.ac.uk/files/2013/01/DataPool_Training_Matrixv1.xlsx) (see Byatt, 2013). Each of the domains was reviewed to see what the researcher needed to understand, what they needed to be able to do and whether it applied to research data. Having identified some areas that were applicable, these were then broken down into the ‘when’ and ‘what’ type of support the researchers might seek and therefore who they might contact for help from within the professional services. The ‘when’ was important as this might influence who they might be in contact with and therefore most likely to approach for assistance. The ResearchData@Soton service, described in the *Research Data Management Planning, Guidance and Support* report (Byatt & White, 2013), had not been launched when this work was done so was not a factor. Drawing on all this information, Harry Gibbs worked toward identifying some of the basic areas that we needed to investigate further with our key groups.

Once the areas were identified our focus turned to how we would find out what individuals in the relevant professionals services (iSolutions, Library and Research and Innovation Services (RIS)) thought they knew about data management. We also wanted to find out whether they felt confident in advising researchers, whether they considered this to be part of their remit and what they felt that they wanted to know about. It was decided that the most appropriate approach would be to use a questionnaire. In discussion with the relevant managers it was thought likely that some staff would have had minimal exposure to the term or concept of “research data management”, while in practice they applied data management principles in their own work. This was taken into account in the development of the questionnaire.

As important as knowing the perceived level of knowledge of the staff was when they didn’t know something, whether they thought they knew who or where to direct researchers. Linked to this was also the question of whether they thought this was something that they ought to know. This would be important to try to find out as it could influence the way that training is presented and help with building links between services where needed.

One further consideration that influenced the development of the survey was that of re-use of the questionnaire at other institutions. As support infrastructure for research data management is a challenge for all higher education institutions, and there was value in collaborating with others in facing them, there would be value in developing a resource that could be used more widely. This would require that it was written in a way that allowed it to be easily adapted for the staff groupings in any institution and wasn’t dictated by the faculty/professional services structure in the University of Southampton. As links already existed with the University of Oxford and they also had a JISCMRD project, it was suggested that the questionnaire could also be trialled with their staff and possibly used to inform some joint training events in the future. With this in mind, the scoping work and early versions of the questionnaire were shared with Meriel Patrick, Analyst and Training Officer, DaMaRO Project, University of Oxford before the survey was finalised. This ensured that with minor changes the questionnaire could also be offered to staff at the University of Oxford and this happened in April 2013.

### The Survey

The University of Southampton provides a survey tool iSurvey (<https://www.isurvey.soton.ac.uk/>), developed in-house and is free to staff and students. It offers many of the basic survey features and as a result it was felt that as the needs of the survey were fairly straightforward and did not warrant the use of more sophisticated packages. A copy of the University of Southampton questionnaire can be found in [Appendix A1](#_JISCDataPool:_Training_Needs)

The survey was divided into 3 sections.

1. Knowledge of research data and its management
2. Directing researchers to the right person
3. About you

The aim of the first section was to explore the perceived level of knowledge or confidence in responding to questions on research data and research data management. In the second section the same topics were offered, but here the emphasis was on whether they felt confident in knowing who to refer a researcher to if they did not consider that it was their job to provide the information. If the respondents saw it as their role they were asked to mark it as “My Role”. The final section collected information about the respondents, their experience and perceived training needs.

In sections 1 and 2 the topics covered were roughly divided into three areas

1. Planning for data management
2. Data Management during the research process
3. Post research data management

The scoping work discussed earlier had identified a fairly wide range of possible topics that could be asked about; this had been reviewed and refined. From the work undertaken by Mark Scott and others (2013) in presenting the *Introducing Research Data* guide we were aware that some of the more basic information had been found the most useful in helping researchers with their work. This influenced the order of some of the topics that we included early in the survey, such as file naming and its role in helping retrieve data, both by the researcher and for a wider audience. The topics that we included in the survey were:

* file naming
* version control
* Intellectual Property (IP), copyright, data protection, and other legal issues
* ethical issues
* re-use of data by funders
* data management plans
* storage requirements and costs
* data security
* metadata – for project and for dissemination
* dissemination
* licensing
* longer-term preservation and archiving and associated costs
* funder requirements on dissemination and sharing

Once the survey had been developed, it required ethics approval before it could be released to the various groups. This process was completed in December 2012.

## The Analysis

The survey was offered to staff at the University of Southampton at the start of 2013. A request to complete the survey using iSurvey was sent out via email to three groups within the University of Southampton – the Library, iSolutions (Computing Services) and Research and Innovation Services in January and early February 2013. The email included a link to the survey and a participant information sheet (see [Appendix A2](#_Participant_Invitation_and)).

The selection of participants was decided by the senior manager of the relevant service, but the recommendation was that the individuals would be in a position where they were in contact with or had a role of supporting researchers. It did not specifically require that support for data management had been identified in their remit.

There were 120 individuals identified and 40 who completed the survey (see [table 1](#Table1)) – 33% response rate. Respondents took on average 15 minutes to finish with a minimum time of just over two minutes and a maximum of just under 42 minutes.

|  |  |  |  |
| --- | --- | --- | --- |
| Professional Service | No. Completed | No. of Recipients |  |
| iSolutions | 8 | 23 |
| Library | 22 | 35 |
| Research and Innovation Services | 9 | 62 |
| Information not provided | 1 |   |
| Total | 40 | 120 |

Table 1 Number of respondents by Professional Service

This is a small sample size and the evaluation and interpretation of results had to be done in this context. However, there is valuable management information in the responses to help inform the focus of future training. The dataset will be added to as the survey is offered to further staff groups at the University of Southampton in the future and staff at the University of Oxford.

The results were analysed using SPSS v20. An initial index of confidence analysis was carried out by Federico De Luca (see separate Appendices B-D). His analysis was done by combining all responses in each of the subsections in sections 1 and 2, so that the minimum, maximum and mean values of confidence were derived. Results were then compared with this index to evaluate the level of confidence.

Table 2. Index of Confidence for Section 1. Planning data management

For example, he looked at all the responses to the section on planning data management - see [table 2](#Table2). The index for this subsection was built in order to go from 0 (minimum confidence) to 36 (maximum confidence) with a mean value of 18. The results for all response showed that for this section more than 75% of respondents had a confidence of 14 or less, and the overall mean was only slightly higher than 10, implying that the confidence of the respondents over the items that define “Planning data management” is quite low.

The figures in [table 3](#Table3) show the frequencies data for each of the questions in the same section that the index of confidence refers to and appears to confirm the overall low confidence in the various aspects of research data management covered in this section.

|  |
| --- |
| **Statistics** |
|   | File naming | Version control | IP & other legal | Ethical issues | Re-use of Data | DMPs |
| N | Valid | 39 | 39 | 40 | 40 | 40 | 40 |
| Missing | 1 | 1 | 0 | 0 | 0 | 0 |
| Mean | 2.54 | 2.62 | 2.93 | 3.30 | 2.90 | 2.13 |
| Median | 2.00 | 2.00 | 3.00 | 3.00 | 3.00 | 2.00 |
| Std. Deviation | 1.847 | 1.830 | 1.607 | 1.786 | 1.646 | 1.453 |
| Minimum | 1 | 1 | 1 | 1 | 1 | 1 |
| Maximum | 7 | 7 | 6 | 6 | 6 | 6 |

Table 3. Planning data management – all responses

As well as analysing all respondents together, De Luca also completed an index of confidence analysis for the Library and for iSolutions/RIS combined. In [figure 1](#Figure1) you can see that there are differences hidden in the results with a wider range in the levels of confidence in the library than in the combined services iSolutions/RIS. The iSolutions/RIS results more closely matched the overall analysis.

 

iSolution/ RIS Combined

Library

Figure 1. Planning data management – Library compared to iSolutions/RIS

A review of the indexes of confidence for the other sub-sections in section 1 and in section 2 of the questionnaire (see separate Appendices B-D) showed that there were differences between the groups and that it would be useful to do further analysis on the individual questions.

This further analysis was carried out by Dorothy Byatt using frequencies and cross-tabulations to review the data. This time analysis was carried out for each question in each section to see if there were any difference between the topics covered and the confidence of staff to respond to them. As well as this analysis being done for all respondents together, it was also carried out for each of the professional services.

Looking at the respondents it was interesting to note that the majority had not been involved in research at advanced postgraduate degree level or as a researcher ([fig. 2](#Figure2)). This is useful in that it can inform the type of information that should be included in training. It is equally useful to know that the staff may not be relying on older practices or presumptions.

Figure 2. Direct Involvement in research

Also from the cross tabular analysis it was noted that those who had been involved in research came from the Library and RIS (see [fig. 3](#Figure3)), but with the low number of responses this may not be a true reflection of all the staff involved.



Figure 3. Direct Involvement in research by Professional Service

### Section 1. Your knowledge of research data and its management

In section 1 the respondents were asked to score themselves in a range where 1 was “Not Confident” to 7 indicating they were “Completely Confident” to give advice on the individual topics mentioned. Most respondents scored their responses in the lowest two categories of confidence in every case in section 1. In order, the areas of most concern to respondents were:

S1.Q3.2: Licensing of the data for re-use ([fig.4](#Figure4))

S1.Q3.4: Costs ([fig.5](#Figure5))

S1.Q1.6: Writing data management plans ([fig.6](#Figure6))

S1.Q3.5: Funder’s requirements for re-use ([fig.7](#Figure7))

S1.Q1.2: Version control ([fig 8](#Figure8))

S1.Q3.3: Long term preservation of data ([fig.9](#Figure9))

S1.Q1.1: File naming ([fig. 10](#Figure10))

S1.Q2.1: Identify data storage requirements ([fig.11](#Figure11))

The following figures show these results split by the Professional Service of the respondents.



Figure 4. Licensing of the data for re-use



Figure 5. Costs



Figure 6. Writing data management plans



Figure 7. Funder’s requirements for re-use



Figure 8. Version control



Figure 9. Long term preservation of data



Figure 10. File naming



Figure 11. Identify data storage requirements

These figures do confirm the results from the index of confidence that respondents’ confidence is skewed towards the ‘not confident’, regardless of the Professional Service. This was not unexpected, but what was of more interest were the small number of ‘6’ and ‘completely confident’ responses given. These may be partly related to the background of individual members of staff. However, it may be as a result of engagement with DataPool activities, in particular through the ResearchData@Soton service that drew in responses from staff in RIS and from iSolutions Business Relationship Managers.

### Section 2. Directing researchers to the right person

This section was a combination of questions. Its aim was to find out whether respondents were confident in referring on a question from a researcher to someone they thought would be the ‘right’ person, organisation, University department or resources to help. This was only where the respondent did not consider that they were the person who should respond because it was part of their work. If this was the case they were asked to say that it was “My Role”. Respondents were also asked whether they thought as part of their work they ‘should know’ who to direct researchers to.

The results show that there is more confidence in knowing who to refer researchers to than providing advice and that there is a wider spread of answers given. Respondents were most confident about ethical and intellectual property issues.

In this section respondents were most uncertain about the correct person to refer the researcher to in the following areas:

S2.Q5.4: Refer costs of data curation ([fig.12](#Figure12))

S2.Q5.1: Refer dissemination and sharing ([fig.13](#Figure12))

S2.Q5.2: Refer Licensing data for re-use ([fig.14](#Figure14))

These are shown in the following figures:



Figure 12. Refer costs of data curation



Figure 13. Refer dissemination and sharing



Figure 14. Refer Licensing data for re-use

Most of the other answers show a pattern like this:

|  |  |
| --- | --- |
| Figure 15. Intellectual property & legal referral  | Figure 16. Funder requirements for data  |

As mentioned earlier respondents were asked to indicate whether they considered an area to be part of their work. In figures 15 and 16 above the “My Role” responses can be seen in the first column (green) on the left. It is interesting that no-one considered the following areas to be their role to give researchers advice on:

S2.Q3.1: Data storage requirements

S2.Q5.2: Refer Licensing data for re-use

S2.Q5.4: Refer costs of data curation

This might have arisen as a result of the sample size, the staff and the services that were asked to participate. However, it suggests that there will need to be a discussion amongst service managers about which service should be the primary provider of support in these areas and how other groups would contribute to or signpost that support. Some agreement would be beneficial before training is offered in these areas.

It was also noted that respondents from both the Library and RIS considered that knowing “Requirements of specific Funders' to make research data available for re-use” and “Writing data management plans for submitting with bids/research proposals to funders” were their roles. It is quite possible that this would be the case where knowledge needs to be in both services as they may interact with researchers in different environments. The experience of the ResearchData@Soton service (Byatt & White) suggests that this is the case; where responding to researchers has required the input from more than one service the majority of time. It will be important to incorporate into the training a discussion on how best these groups can continue to interact to ensure that they provide a consistency in the responses given. It also illustrates why there will be value in have joint training events.

Analysis for all respondents to the question about whether the respondent ‘should know’ who to refer the researcher to does not bring out any clear picture, although in 10 out of the 15 areas the number of respondents who say they ‘should know’ is equal or higher than those who didn’t. When broken down into the Professional Services it does appear to show that there are some differences between services in what they consider they should know. For example there are some topics where a professional service response is opposite to the overall response. This is illustrated in the question on funder requirements for re-use. [Figure 17](#Figure17) shows that the majority say they ‘should know’, but [figure 18](#Figure18) shows the majority of iSolutions respondents did not say so.

Figure 17. Figure 18.

### Section 3. About the respondents

The beginning of the analysis discussed the number of respondents by the Professional Services and their research experience.

In this section respondents were also asked about whether they had had any requests for help with these aspects of research data management within the last 6 months and if so how many times. At the time of the survey the University Policy had been in place for nearly ten months and the Research Councils’ policies on research data for longer. It is interesting that the majority of respondents had not had any such requests (see [fig.19](#Figure19)), but this reflects the high number of library respondents who may not yet have become directly involved in research data management support. However, only one of the RIS respondents had not had a request in this area (see [fig.20](#Figure20)) and RIS as the main research support service, not unexpectedly, received the highest number of requests.

Figure 19.  Figure 20.

At the end of each section respondents were asked if there were any particular areas that they wanted training. This was to capture a free response to the various topics with the opportunity to describe what they wanted in an undefined way. In the final section respondents were asked to select from a prescribed list that grouped some areas together such as file naming and versioning under file management. The results for the final section are shown in [table 4](#Figure4) and show that data management plans was the area most requested, closely followed by security and storage. The totals were close for all those selected for the data re-use issues.

|  |  |
| --- | --- |
| **Most essential area of training is** |  |
| Data management plans | 9 |
| Legal issues and research data | 7 |
| Making data available for re-use | 2 |
| Security and storage of data | 8 |
| Describing your data | 6 |
| Funder requirements for research data | 6 |
| Other, please specify | 1 |
| Total | 39 |

Table 4 Most essential training requested

Looking at the breakdown by Professional Service (see [fig.21](#Figure21)) we see that the Library has the widest range of training requests, while iSolutions and RIS concentrated on more focussed areas. One person did request training in another area and this was for something on research data management within a specific discipline.

Figure 21. Training requested by Professional Service

Three areas were not selected from the list offered and these included file management - naming and versioning, data management costs, ethical issues and data. Interestingly file naming and version control was an area respondents said that they were not confident in providing advice. Respondents had, however included them in the training listed at the end of section 1 and section 2. [Figure 21](#Figure21) shows the results with these added in. This confirms the five key areas to concentrate on in training.

Figure 22. All training requested

## Case Study: Introductory training -the library

Using the early results of the analysis pilot training was held for library staff. This consisted of a one hour ‘briefing’ and a 90 minute workshop as a linked event. The briefing covered the background to the increased emphasis on research data, funder requirements, the University of Southampton policy, the data life cycle as well as some of the practical aspects that need to be considered by researchers. This incorporated information from *Introducing Research Data* (Scott, 2012); the services offered via the DCC DMPonline tool and described the ePrints data deposit option[[1]](#footnote-1).

In response to the request for training in data management planning the workshop focussed on this. The curriculum on research data management, released by the Information School, University of Sheffield (2013) as part of the JISC RDMRose project, was reviewed and informed our approach to the workshop. It was based around a real research proposal. The first exercise asked participants to identify the 5 most important things to be included in the data management plan for that proposal. The second exercise covered the type of data that might be created by the project, what issues this might raise and finally what they might ask the researcher to check that these areas had been considered. The feedback on these sessions was very positive (see [Appendix A3](#_Pilot_training_event)) and participants agreed at the end that they would be able to begin the process of helping a researcher with a data management plan.

Although this was only open to library staff it was not specific to library roles and would transfer to the other staff groups. This could also be used to facilitate joint events as the workshop would certainly have allowed for the different groups to share their expertise. Discussions with service managers in response to the survey will include this offer.

## Summary and the way forward

Our aim was to create a survey that could be used by different groups within an institution; that it would draw out the knowledge held within staff and their awareness of support within the organisation; and that it could be used in more than one institution. By reviewing the different stages in the scoping process referred to earlier, letting this inform the key aspect to include in the survey and drawing in advice from the Library, Social Sciences department and the University of Oxford on the development of the questionnaire, the survey developed is one that would meet this need.

The results of the analysis for the University of Southampton did show what we were expecting. It shows that although the majority may not be confident there are a few who do have confidence in some of the key areas. It also showed that there was an awareness of support elsewhere. It might have been useful to ask who they thought this would be, but this would have added a further layer of complexity to the survey and on balance was not thought appropriate at this point. It would have allowed us to identify whether this was through formal or informal networks. However, it will be important to draw on the expertise identified and build on support network in the future training events. This could easily be incorporated into the structure of the briefing/workshop training outlined above.

The design of the questionnaire was strongly influenced by the desire that it could be used elsewhere with minimal alteration. This happened in April 2013 when the survey was opened up to staff at the University of Oxford. The survey was administered by the University of Southampton using iSurvey. The main changes that were required were the staff groupings and this was a simple process in the iSurvey platform. It was also possible to add a question in section 1 and 2 relating specifically to the University of Oxford research data management policy. The results of this and a comparison with Southampton results will be reported separately. There did not seem to be any problems in rolling this out in this manner.

At the University of Southampton the next steps will be to fully discuss the findings with the managers of the relevant services to see how training and support can be developed for the relevant staff. Also there were some groups of staff that were not included in the survey for a number of reasons and it will be important to engage with them so that we can draw on their knowledge, particularly in the area of Finance to help with advice on costing.

It is a recommendation that in the future any training provided includes staff from all the professional services. This will facilitate the raising of awareness of what other Professional Services do, to recognise their expertise and to develop a co-ordinated network of support for the benefit of the researcher. Evidence that this network is beginning to develop can be seen by the services involved who helped to respond to requests sent to the ResearchData@Soton service, outlined by Byatt & White (2013, table 2).

Following the involvement of the University of Oxford there will be benefits from developing joint training events so that support comes from a background of broader understanding of the process involved in research data management including how others approach this area. Equally important will be to draw in expertise from the Higher Education community, such as the Digital Curation Centre, in relation to data curation and the longer term preservation aspects.

## References

Brewerton, Antony (2012) Re-skilling for research: Investigating the needs of researchers and how library staff can best support them *New Review of Academic Librarianship* Vol.18 No.1, pp. 96-110 <http://dx.doi.org/10.1080/13614533.2012.665718>

Brown, Mark, Parchment, Oz and White, Wendy (2011) Institutional data management blueprint. Southampton, GB, University of Southampton, 21pp <http://eprints.soton.ac.uk/id/eprint/196241>

Byatt, Dorothy (2013) *Mapping training needs for the support team* DataPool blog <http://datapool.soton.ac.uk/2013/01/21/mapping-training-needs-for-the-support-team/> Training Matrix <http://datapool.soton.ac.uk/files/2013/01/DataPool_Training_Matrixv1.xlsx>

Byatt, Dorothy and White, Wendy (2013) *Research Data Management Planning, Guidance and Support: a DataPool Project report* Southampton, GB: University of Southampton <http://eprints.soton.ac.uk/id/eprint/351027>

Byatt, Dorothy, Scott, Mark, Beale, Gareth, Cox, Simon J, White, Wendy (2013) *Developing Researcher Skills in Research Data Management: Training for the future - a DataPool project report.* Southampton, GB: University of Southampton <http://eprints.soton.ac.uk/id/eprint/351026>

Digital Curation Centre (2011) *Research Data Management Skills Support Initiative* (DaMSSI) <http://www.dcc.ac.uk/training/damssi>

Digital Curation Centre (2013) *Data Management Skills and Support Initiative – Assessment, Benchmarking and Classification* (DaMSSI-ABC)

Research Councils UK (2011) *Common principles on data* policy [web page] <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx> (accessed 21/05/2013)

Research Councils UK (2012) *RCUK Welcomes Royal Society “Science as an Open Enterprise” report* [press release] <http://www.rcuk.ac.uk/media/news/2012news/Pages/120621.aspx>

Royal Society (2012) *Science as an open enterprise* The Royal Society Science Policy Centre report 02/12. London: Royal Society <http://royalsociety.org/policy/projects/science-public-enterprise/report/>

Scott, Mark, Boardman, Richard P., Reed, Philippa A.S. and Cox, Simon J. (2012) Introducing research data. Southampton, GB: University of Southampton, 29pp <http://eprints.soton.ac.uk/id/eprint/338816>

University of Sheffield Information School (2013) *RDMRose Learning Materials* Sheffield: University of Sheffield <http://rdmrose.group.shef.ac.uk/?page_id=10>

Vitae (2010) *Vitae Researcher Development Framework* Careers Research and Advisory Centre (CRAC) Limited <http://www.vitae.ac.uk/rdf>

Vitae; Research Information Network; SCONUL (2012) *Information Literacy Lens on the Researcher Development Framework* Careers Research and Advisory Centre (CRAC) Limited <http://www.vitae.ac.uk/researchers/1272-521621/Information-Literacy-Lens-on-the-Researcher-Development-Framework.html>

White, Wendy and Brown, Mark (2013) *DataPool: Engaging with our Research Data Management Policy.* Southampton, GB, University of Southampton, 6pp. <http://eprints.soton.ac.uk/id/eprint/351945>

## Acknowledgements

We would like to acknowledge specific thanks for the help and support that we received in the development and analysis of the survey and questionnaire. Advice was received from a number of people and in particular from:

**Harry Gibbs**, University of Southampton Library

**Teresa McGowan**, Research Manager, Social Sciences, Faculty of Social and Human Sciences, University of Southampton

**Laura Molloy**, Research Officer, HATII, University of Glasgow/ JISC DaMSSI-ABC project

**Meriel Patrick**, Analyst and Training Officer, DaMaRO Project, University of Oxford

**Sally Rumsey**, DaMaRO Project, Bodleian Libraries, University of Oxford

## Copyright

© 2013 University of Southampton

This work is licensed under a [Creative Commons Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/deed.en_GB).

You are free to re-use part or all of this work elsewhere, with or without modification. In order to comply with the attribution requirements of the Creative Commons license (CC-BY), we request that you cite:

Byatt, Dorothy, De Luca, Federico, Gibbs, Harry, Patrick, Meriel, Rumsey, Sally and White, Wendy (2013) *Supporting researchers with their research data management: Professional service training requirements - a DataPool project report.* Southampton, GB, University of Southampton

# Appendix A

## JISCDataPool: Training Needs

Traditionally data created through research has been analysed and only a portion has been made available through papers, conferences, reports, articles etc. However the importance of source data has been recognised by the UK Government, Research Funding Councils and others. Policies have been developed requiring researchers to actively manage all their data, from concept to conclusion, to ensure that any significant data can be held safely and securely for future re-use. As a result, institutions need to develop appropriate systems, support and guidance to assist researchers in this task. This support starts with the research proposal and ends with the storage of data in an appropriate repository.

It is recognised that staff that support researchers, particularly in roles like yours, will be asked for advice. In order to help you and others working in this area, we would like to identify your training needs and develop appropriate guidance. We really need your help with this.

The questionnaire is confidential and anonymous. Your managers will not see your individual responses. It should only take about 15 minutes to complete, but it will be invaluable in identifying both strengths and gaps in knowledge.

This work is being funded by the JISCMRD DataPool Project and a report will be written as part of the project. If you have any questions about this project or questionnaire please contact Dorothy Byatt (drb@soton.a.uk) or Wendy White (whw@soton.ac.uk).

The questionnaire is divided into three sections. In the first we need an honest assessment of how confident you are in each of the main areas of support. In the second we need to understand how confident you are that you would know where to refer an enquirer if you were not the right person to help them. Finally we need a few facts about you.

**Participant Consent form**

Please read the following statements carefully before agreeing to take part in this study

* I have read and understood the participant information sheet (attached to the email in which you received the link to this questionnaire)
* I agree to take part in this research project and agree for my data to be used for the purpose of this study
* I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected
* I am happy for my anonymised responses to be presented in summary form along with the rest of the participants and made publically available

#### Section 1. Your knowledge of research data and its management

For each question please score your answer from 1 to 7 where 7 means that you are completely confident in your knowledge to handle an enquiry on this topic and 1 means that you feel you are not confident and need more knowledge in this area.

##### Planning for data management

###### Question 1.1

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   | Not confident | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Completely confident |
| Advise on the use of file naming to assist with file management and retrieval |   |  |  |  |  |  |  |  |   |
| Advise on version control of files so that it forms good practice |   |  |  |  |  |  |  |  |   |
| Assist in identifying Intellectual Property (IP), copyright, data protection, and other legal issues relating to research data |   |  |  |  |  |  |  |  |   |
| Be able to help identify ethical issues relating to research data, for example, issues of confidentiality |   |  |  |  |  |  |  |  |   |
| Advise on the requirements of specific funders to make research data available for re-use |   |  |  |  |  |  |  |  |   |
| Advise on writing data management plans for submitting with bids/research proposals to specific funders |   |  |  |  |  |  |  |  |   |

##### Data Management during the research process

###### Question 1.2

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   | Not confident | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Completely Confident |
| Help identify the data storage requirements - size, type - during the lifetime of the project |   |  |  |  |  |  |  |  |   |
| Raise awareness of data security issues during the lifetime of the project |   |  |  |  |  |  |  |  |   |
| Advise on how to describe the data so that others on the project can find the data, i.e. adding basic working metadata |   |  |  |  |  |  |  |  |   |
| Advise on describing the data so that others in the future will be able to find and safely re-use the data, i.e. adding public metadata |   |  |  |  |  |  |  |  |   |

##### Post research data management

###### Question 1.3

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   | Not confident | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Completely confident |
| Advise on options for the dissemination and sharing of data |   |  |  |  |  |  |  |  |   |
| Advise on issues associated with the licensing of the data for re-use |   |  |  |  |  |  |  |  |   |
| Advise on the issues associated with the longer-term preservation and archiving of data |   |  |  |  |  |  |  |  |   |
| Advise on the costs associated with the creation of data during the project including storage, dissemination and preservation |   |  |  |  |  |  |  |  |   |
| Advise on specific funders requirements to allow the exploitation and re-use of data |   |  |  |  |  |  |  |  |   |

###### Question 1.4

Some staff have told us that they feel more confident about some aspects than others. If there is one or two areas in this section where you feel you particularly need further training, please list them below.

|  |
| --- |
|   |

 Thank you for completing this section

#### Section 2. Directing researchers to the right person

Due to the nature of our work there are aspects of research data management that we may be asked about but are not responsible for supporting. In this section we would like to find out how confident you are in being able to refer the researcher to the right person, organisation, University department or resources to help?

In this section the topics are the same, but we would like you to score your answer from 1 to 7 where 7 means that you are completely confident in your ability to correctly refer an enquiry on this topic and 1 means that you feel you have low confidence that you could refer an enquiry to the right place.

If you consider that answering a question on the topics below is part of your remit, please tick 'My Role' and move to the next topics.

##### Planning for data management

###### Question 2.1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   | Low confidence | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Completely confident | My Role |
| Use of file naming to assist file management and retrieval |   |  |  |  |  |  |  |  |   |  |
| Version control of files so that it forms good practice |   |  |  |  |  |  |  |  |   |  |
| Identifying Intellectual Property (IP), copyright, data protection, and other legal issues relating to research data |   |  |  |  |  |  |  |  |   |  |
| Identifying ethical issues relating to research data, for example, issues of confidentiality |   |  |  |  |  |  |  |  |   |  |
| Requirements of specific Funders' to make research data available for re-use |   |  |  |  |  |  |  |  |   |  |
| Writing data management plans for submitting with bids/research proposals to funders |   |  |  |  |  |  |  |  |   |  |

###### Question 2.2

The topics listed cover a wide area with some more related to your work than others so you may feel that you **should know** who to refer questions to when asked. For each of the topics where you think that is the case, please tick the 'Should know' box provided, whether or not you are confident at present. Tick all that apply.

|  |  |
| --- | --- |
|   | Should know |
| Use of file naming to assist with file management and retrieval |  |
| Version control of files so that it forms good practice |  |
| Identifying Intellectual Property (IP), copyright, data protection, and other legal issues relating to research data |  |
| Identifying ethical issues relating to research data, for example, issues of confidentiality |  |
| Requirements of specific funders to make research data available for re-use |  |
| Writing data management plans for submitting with bids/research proposals to specific funders |  |

##### Data management during the research process

###### Question 2.3

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   | Low confidence | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Completely confident | My Role |
| Identifying the data storage requirements - size, type - during the lifetime of the project |   |  |  |  |  |  |  |  |   |  |
| Awareness of data security issues during the lifetime of the project |   |  |  |  |  |  |  |  |   |  |
| How to describe the data so that others on the project can find the data, i.e. adding basic working metadata |   |  |  |  |  |  |  |  |   |  |
| How to describe the data so that others in the future are able to find and safely re-use the data, i.e. adding public metadata |   |  |  |  |  |  |  |  |   |  |

###### Question 2.4

The topics listed cover a wide area with some more related to your work than others so you may feel that you **should know** who to refer questions to when asked. For each of the topics where you think that is the case, please tick the 'Should know' box provided, whether or not you are confident at present. Tick all that apply.

|  |  |
| --- | --- |
|  | Should know |
| Identifying the data storage requirements - size, type - during the lifetime of the project |  |
| Awareness of data security issues during the lifetime of the project |  |
| How to describe the data so that others on the project can find the data, i.e. adding basic working metadata |  |
| How to describe the data so that others in the future will be able to find and safely re-use the data, i.e. adding public metadata |  |

##### Post research data management

###### Question 2.5

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   | Low confidence | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Completely confident | My Role |
| Options for the dissemination and sharing of data |   |  |  |  |  |  |  |  |   |  |
| Issues associated with the licensing of the data for re-use |   |  |  |  |  |  |  |  |   |  |
| Issues associated with the longer-term preservation and archiving of data |   |  |  |  |  |  |  |  |   |  |
| Identifying the costs associated with the creation of data during the project including storage, dissemination and preservation |   |  |  |  |  |  |  |  |   |  |
| Requirements of specific funders to allow the exploitation and re-use of data |   |  |  |  |  |  |  |  |   |  |

###### Question 2.6

The topics listed cover a wide area with some more related to your work than others so you may feel that you **should know** who to refer questions to when asked. For each of the topics where you think that is the case, please tick the 'Should know' box provided, whether or not you are confident at present. Tick all that apply.

|  |  |
| --- | --- |
|   | Should know |
| Options for the dissemination and sharing of data |  |
| Issues associated with the licensing of the data for re-use |  |
| Issues associated with the longer-term preservation and archiving of data |  |
| Identifying the costs associated with the creation of data during the project including storage, dissemination and preservation |  |
| Requirements of specific funders to allow the exploitation and re-use of data |  |

###### Question 2.7

As each of our roles differ there may be one or two areas in the topics mentioned above where you feel you particularly need further guidance on where to direct people, please list them below.

|  |
| --- |
|   |

  Brilliant! Just one short section left.

#### Section 3. About you

In order to understand where to target advice and assistance we really need to know a bit about you. All results are confidential and we do not ask for your name.

###### Question 3.1

I am a member of staff working in

|  |  |
| --- | --- |
|  | Finance  |
|  | iSolutions  |
|  | Library  |
|  | Research and Innovation Services  |
|  | Faculty Administration - Graduate School  |
|  | Doctoral Training Centre  |
|  | Governance  |
|  | Legal Services  |
|  | Other, please specify  |

Question 3.1b [completed if Other, please specify selected]

I work in ...

*(Please give the area that you are working in, not necessarily your job title)*

|  |
| --- |
|   |

###### Question 3.2

What is your role in the University where you are supporting researchers?

*If your job is divided between roles, please list the one(s) in which you support researchers.*

|  |
| --- |
|   |

###### Question 3.3

Have you been directly involved in carrying out research as a researcher or as part of an advanced postgraduate research degree?

Please tick all that apply

|  |  |
| --- | --- |
|   | Tick |
| Yes, as a researcher |  |
| Yes, as part of an advanced research degree |  |
| No |  |

###### Question 3.4

Have you been asked for help related to research data management in the last 6 months?

|  |  |
| --- | --- |
|  | Yes, once only  |
|  | Yes, between 2-5 times  |
|  | Yes, more than 5 times  |
|  | No  |

###### Question 3.4b

Why do you think that is?

|  |
| --- |
|   |

###### Question 3.5

If you had to choose one area of training as being the most essential what would it be? Please select on from the drop down list.

|  |  |
| --- | --- |
|  | File management - naming and versioning  |
|  | Data Management Plans  |
|  | Legal issues and research data  |
|  | Ethical issues and data |
|  | Making data available for re-use |
|  | Security and storage of data |
|  | Describing your data |
|  | Funder requirements for research data |
|  | Data management costs |
|  | Other, please specify |

If you can, please briefly explain why.

|  |
| --- |
|   |

Question 3.5b [completed if Other please specify is selected]

If you can, please briefly explain why.

|  |
| --- |
|   |

###### Question 3.5a

Please say what other training you would like and if you can, briefly explain why.

|  |
| --- |
|   |

###### Question 3.6

If you have any comments about research data management please use the space below.

|  |
| --- |
|   |

Thank you so much for taking the time to complete this questionnaire, we know that you are very busy. Your responses will be used to inform our development of guidance and training and your help will be invaluable.

If you would like to contacts us about the project or topics mentioned, please email data@soton.ac.uk

## Participant Invitation and Information

##### Email content inviting participation

I am Dorothy Byatt, Project Manager, JISC DataPool and Academic Liaison Librarian, University of Southampton. You have been sent this email as we consider that the area and nature of your work brings you into contact with researchers who may seek your advice on data management planning, data handling, storage and security or ask you where they can get advice. I am requesting your participation to help us identify the training and guidance we need to provide you with in support of this role. The questionnaire should last approximately 15 minutes.

Personal information will not be released or viewed by anyone other than researchers involved in this project. By taking part you will be helping us to develop the training programme and guidance for Professional staff, and thereby the quality of the support offered to researchers

All responses are treated as confidential, and in no case will responses from individual participants be identified. Rather, all data will be pooled and published in aggregate form only.

The survey is available from <https://www.isurvey.soton.ac.uk/6254>

Please read the attached participant information sheet for further information and if you have any questions please contact me, Dorothy Byatt, Project Manager (drb@soton.ac.uk) or Wendy White, Principal Investigator, (whw@soton.ac.uk)

##### Participant Information Sheet

**Study Title**: DataPool – Training Needs

**Researcher**: DataPool **Ethics number**: 4560

**Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.**

**What is the research about?**

This research is being undertaken as part of the JISC funded DataPool project. DataPool is about embedding good data management practices into the infrastructure and culture of the University. To do this it is important to develop the skills and expertise of those staff that support researchers with data management. This questionnaire is designed to find out what expertise is currently available and what areas require staff development.

**Why have I been chosen?**

You have been chosen as we have identified that your area of work brings you into contact with researchers when they may need advice or assistance with some aspect of research data management, for example, writing a proposal, looking for storage, advice on dissemination tools.

**What will happen to me if I take part?**

The questionnaire should take about 10-15 minutes of your time. Your responses will be analysed along with other participants and used to develop appropriate guidance material and training.

**Are there any benefits in my taking part?**

You will be helping to inform the training that will be offered to help develop your skills in areas related to your work.

**Are there any risks involved?**

There are no risks

**Will my participation be confidential?**

Although you have been invited to complete the questionnaire it is confidential, and will conform to the Data Protection Act/University policy. We do not ask your name, but we do ask for your role. Questionnaire data will be made openly available as part of the research output but information relating to your area of work will be generalised. Unlinked anonymity of participants cannot be assured.

**What happens if I change my mind?**

You may withdraw at any time without you legal rights being affected.

**What happens if something goes wrong?**

In the unlikely case of concern or complaint please contact Dr Martina Prude, Head of Research Governance (02380 595058, mad4@soton.ac.uk).

**Where can I get more information?**

Contact Wendy White (whw@soton.ac.uk) DataPool Principal Investigator or email data@soton.ac.uk if you require more information about this project.

## Pilot training event – Library – evaluations

### Professions Services – Library Briefing

|  |  |
| --- | --- |
| **Course:** | Research Data Briefing - Library |
|  |  |  |  |  |
| **Forms Returned** | 13 |  |  |  |
| Number attended  | 19 |  |  |  |
| **Presenter** | Harry Gibb; Dorothy Byatt; |  |
|  |  |  |  |  |
| The presentation and facilitation skills of the event leader(s) | **[Score]** | 4.2 |   |  |
| Overall rating of the event | **[Score]** | 4.2 |  |  |
| The administrative arrangements prior to the event | **[Score]** | 4.4 |  |  |
| The content of the event was as advertised | **[Score]** | 4.5 |  |  |
| The event addressed my needs | **[Score]** | 4.2 |  |  |
| Would you recommend this event to a colleague/fellow student? | **Yes** | 61.5% |  |  |
|   | **Maybe** | 38.5% |  |  |
|   | **No** | 0.0% |  |  |
| [**Score** 1= Very Unlikely / Very Poor; 5 = Very Likely / Very Good] |
|  |  |  |  |  |
|

|  |
| --- |
|  |

 |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

### Professional Services – Library Workshop

|  |  |
| --- | --- |
| **Course:** | Research Data Management Planning Workshop |
|  |  |  |  |  |
| **Forms Returned** | 16 |  |  |  |
| Number attended  | 19 |  |  |  |
| **Presenter** | Harry Gibb; Dorothy Byatt; |  |
|  |  |  |  |  |
| The presentation and facilitation skills of the event leader(s) | **[Score]** | 4.9 |   |  |
| Overall rating of the event | **[Score]** | 4.8 |  |  |
| The administrative arrangements prior to the event | **[Score]** | 4.8 |  |  |
| The content of the event was as advertised | **[Score]** | 4.8 |  |  |
| The event addressed my needs | **[Score]** | 4.7 |  |  |
| Would you recommend this event to a colleague/fellow student? | **Yes** | 93.8% |  |  |
|   | **Maybe** | 6.3% |  |  |
|   | **No** | 0.0% |  |  |
| [**Score** 1= Very Unlikely / Very Poor; 5 = Very Likely / Very Good] |
|  |  |  |  |  |
|  |  |  |  |  |
|

|  |
| --- |
|  |

 |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

1. This option is based the work carried out by the University of Essex as part of the Data@Essex project. [↑](#footnote-ref-1)