# Research Data Management Planning, Guidance and Support: a DataPool Project report

**Dorothy Byatt,** JISC DataPool Project, University of Southampton Library, University of Southampton, UK; **Wendy White**, Principal Investigator, JISC DataPool Project, University of Southampton Library, University of Southampton

## 1. Introduction

This report will review the development of research data management support in the University of Southampton following the approval of its research data management policy in February 2012. Wendy White (2013) in her report *DataPool: Engaging with our Research Data Management* Policy discusses the rationale and approach to the development of the policy. This report will look at the development of the research data management web pages, including the supporting policy guidance, and then will focus on the ResearchData@soton email, phone and desk side service launched to provide research data support to the University.

The University Research Data Management Policy was passed by Senate in February 2012 and was subsequently added to the official University Calendar [http://www.calendar.soton.ac.uk/sectionIV/research-data-management.html]. Linked to the release of the policy was the launch of a suite of web pages containing guidance on a variety of key areas for the management of research data including those specified in the policy. It was hosted on the Library Research website [http://www.southampton.ac.uk/library/research/researchdata/]. Some of the topics covered include Data Management Planning; research data storage options; managing and sharing data; restricting access to data; retention periods and secure destruction of electronic data. Review of how the guidance was developed is described in section 3.

At the same time as the website was released the ResearchData@Soton service was launched, offering assistance to researchers using a variety of response methods, including email, direct conversation and desk side. This service focusses on the support of research data management and curation throughout the research lifecycle and is co-ordinated through an email account supported by a network of staff from a variety of areas. This has already proved to be an important offer not just to the researchers themselves, but also to a number of different professional services in helping to co-ordinate responses to the researcher and will be dealt with in more detail in the following section.

## 2. ResearchData@soton Service

The library began its ResearchData@Soton service offering support for research data management, including desk side sessions, at the same time as the release of the website and of the University policy. The University already had an ethos of Deskside coaching and training to staff and researchers. The University IT Professional service, iSolutions offers a service that covers basic IT software and teaching resources and the Library offers support on information management, including bibliometrics, open access, e-theses, bibliographic databases and reference manager software. Although it might have been feasible to add a service for research data management to these existing ones it was believed that a more bespoke offer was required. The primary reason for this was the expectation that the responses would not be internally focussed, i.e. within one service, but look to the network of services that support research within the institution. This network includes Research and Innovation Services (RIS), responsible for support for research proposal processes including Intellectual Property rights; iSolutions, responsible for the IT/data storage; Finance, responsible for costing the proposals; Governance, whose work includes ethics approval; and the Library, with responsibilities in supporting researcher information needs and management of the Institutional Repository ePrints Soton. The service would offer support, using the most appropriate method of response whether email, Skype or at the desk side of researchers, drawing in those individuals as necessary to address the issues at hand. The aim was to support the researcher where and when it was needed and to involve the most appropriate person who might not be a member of library staff.

The service commenced in early July, and by the end of the month we had made contact with two researchers who were working on grant applications. The first was for an industrial CASE award from the Medical Research Council (MRC) and came in via email with sufficient helpful information for work to begin. Turning to the guidance pages they immediately proved themselves by providing quick links to relevant support and guidance. A phone call and several emails led to the researcher being happy with what he had written in his data management plan, now considered it ready for submission, and that was the end of our first request. Overall it took around 3 hours and had a short turnaround time as the bid was due at the beginning of August. This was an example of the Library being able to direct the researcher to the appropriate information, to raise his awareness of the issues and enable him to complete the necessary work within a tight timescale. Although the first, it has proved invaluable in developing our skills and knowledge of the MRC policy on data (MRC, 2011) and data management plan (MRC, 2012) as a significant number of our enquiries so far have been for support for MRC grants proposals for postgraduate fellowships and research projects.

The second request for help came through academic contact and was a much more complex one. It demonstrated the need for the network of support that we had envisaged. It was for an AHRC proposal that needed to agree data storage and support to host data visualisation of heritage material within the Institution. It involved experts from iSolutions, the Library and the Academic community to help identify the exact requirements to ensure that the proposal had included the resources required. Support was provided through emails and a face-to-face meeting of the group. During this it became clear, from the conversation between the Academics and iSolutions, that in order to develop the technical aspects fully it would be important to have a researcher with the necessary skill on the team. This had not been identified previously and might not have been had the service not been available.

These two requests turned out to be fairly representative of those that followed, some straightforward needing help with identifying the topics to be covered and some were more complex, needing the involvement of a number of different individuals. You can see a summary of some of the key data relating to the service in Table 1. We consider the figure for the number of enquiries to under reflect the actual number of research data management requests for support in the institution.

|  |  |
| --- | --- |
| Number of enquiries | 23 |
| Average length of time to complete enquiry | 3 hours 25 minutes |
| Range of time | 5 minutes to 10 hours |
| Number of Funders | 7 |
| Number asking for Data Management Planning advice | 16 |
| Number asking for Metadata advice | 2 |
| Number asking for storage advice | 1 |
| Number asking for other assistance | 4 |
| Number that required follow-up | 6 |
| Number that were answered in a single enquiry conversation, including phone conversations, emails, interviews | 15 |
| Number that involve one or more services in responding to the request | 13 |
| Source of Enquiry | |
| Direct via email | 13 |
| Research & Innovation | 2 |
| iSolutions | 4 |
| Academic | 3 |
| Google | 1 |

**Table.1 ResearchData@Soton key data**

Existing routes of support from Professional Services staff and Research Groups within Faculties have continued, but cannot be quantified. The networks of expertise will be enlarged to draw these groups in as is appropriate. Currently the average time taken to answer these requests is around 3.5 hours and the longest covering approximately 10 hours. As the skills and experience of those supporting this service develop, and the training identified in the survey of Professional Services (Byatt, 2013) has been provided, it is likely that the average time would reduce. Further resources and training for researchers will be developed, building on the work of Mark Scott et al (2012) and the DataPool Project (Byatt et.al, 2013), and in response to answer frequently answered questions and common issues. This will increase the existing level of expertise in the researcher community. However, there is a significant staff commitment required to assist with the more complex proposals, this is less likely to reduce over time and is an area that will need to be addressed as research data management support continues to be embedded in the Institution.

### Selected Data Enquiries

The following table (table 2.) is a representative selection of the requests for help illustrating the topics raised and resources used to respond to the requests. We received enquiries for the majority of RCUK funder but not NERC or ESPRC during the project. There is a close relationship between the University and the NERC National Oceanographic Centre Southampton so support for the research data management is mainly directed towards the British Oceanographic Data Centre. Although EPSRC require data to be managed and shared, there is no requirement for a data management plan, having placed the responsibility on the institution to have the necessary policy and processes in place. As many of our contacts have arisen out of a need for a data management plan it is not unexpected that EPSRC is not represented at this stage.

|  |  |  |
| --- | --- | --- |
| Topic | Responses | Services involved |
| Research data management for AHRC | Storage options, ePrints Soton  Institutionally hosted data visualisation,  Longer term storage of heritage material,  Working with third party harvesters to maximise impact of material,  Advice on metadata requirements and harvesting standards.  Advise on possible CO-I input for bid and review draft technical summary | Library  Academic  iSolutions |
| Data Management Plan for MRC | Advice on establishing good practice  Identifying and safe handling of sensitive data  Storage and equipment needs  Metadata  Guidance pages; MRC Template and guidance; DMP Online | Library  iSolutions |

|  |  |  |
| --- | --- | --- |
| Topic | Responses | Services involved |
| Data Management Plan for ESRC | Review and comment on draft  Advice on interpreting ESRC guidance  Data storage, ePrints Soton  Sharing and security  Advice on appropriate metadata standards  Guidance pages and University policy; DCC DMP Online | Library |
| Project database and website, AHRC proposal | How to capture data and back-up data when collecting data “in the field”  Handling of different forms of raw data  Creation of database  Guidance on new summary and plan requirements  AHRC Guidance; DMP Online | Library  iSolutions |
| Data management for educational research | Advice on implications of BERA guidelines  Secure storage of data  Guidance and University policy; iSolutions | Library |
| Grant application DMP for MRC | Identified potential for sensitive data and who might advise on best practice in handling such data  Suitable file formats  Guidance on what was meant by ‘type’ of data in MRC template  MRC guidance; DMP Online | Library  iSolutions |
| Project Grant AHRC | Creation of website and database  Equipment and storage requirements | Library  iSolutions  Academic |
| Data Sharing for a continuation project | Data sharing requirements;  Issues associated with sharing or not sharing data  Guidance pages and University policy; DMP Online | RIS  Library  Academic |
| Data sharing for follow-on project | Ethical and sensitive data issues  NCRI website; CRUK guidance | Library |
| Identifying what is Research Data | Assistance in identifying data in a proposal that required supporting documentation  Advice on choice of open formats  Longer-term governance on data with long life span  AHRC Guidance; DMP Online | Library  iSolutions  Academic |

|  |  |  |
| --- | --- | --- |
| Topic | Responses | Services involved |
| Adding Metadata to multiple files | Advice on adding metadata to MP3 files  Guidance from iSolutions | Library  iSolutions |
| What is significant research data; what needs to be kept | Advice on reasons behind data policy  Academic integrity  Future re-use | Library |
| Legal and other requirements for storage and sharing of human related data for EU | Advised on location of EU guidance for project reporting and communication  Directed to contact ethics committee | Library  RIS  iSolutions |

**Table 2. Selected Data Enquiries**

Our approach to responding to these enquires is to engage in a conversation with the researcher to help understand the nature of their research proposal. Often this would be by asking questions about what they currently do, what kind of data they will create and how they expect to disseminate it. From this links can be made to the funder requirements and guidance. Although there is often good guidance provided by the funder it is by necessity generic and being able to help translate it into practice and offer suggestions has been well received. Links will also be made with other Professional Services so that the researcher can bring all the necessary information together for inclusion in the plan.

In preparing to respond to some of the data management plan requests above, it has been important to review guidance and documentation available for the relevant funder. So far in the majority of our enquiries the researcher provided some basic information about who they were applying to and why, but it was always useful to clarify exactly which funder, the bidding call and when they need to submit. The DCC *Overview of funders' data policies* page (<http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>) proved to be a useful starting point although it is still important to look at guidance for the specific proposal. Also the DCCs DMP online tool (<https://dmponline.dcc.ac.uk/>) has been helpful in us to understand the funder requirements in more detail.

Some requests do ask for stock phrases or paragraphs and generally this is not the approach we would take. We intend to continue to offer help so that research data management plans and sharing plans are written in a way that demonstrates the reflection and planning for the full life cycle of a project and avoids the ‘cut and paste’ template approach except where the funder expects this type of information. For example, the MRC asks for links to specific policies and this can be added to a University MRC template. However, the guidance pages do provide some information that can be used in this way and some enquiries can be answered by directing researchers to them.

## 3. Developing the Guidance and Web Site

The development of the guidance and the support for research data management has taken a similar approach to that of the development of the University Policy (White, 2013) in that it was an iterative process and that it involved collaboration with academics, researchers and professional services staff to develop the content. This continues to be the case and additional guidance has already been added to the web pages since the initial launch. Existing guidance has also been improved in response to feedback and to new developments within the research data management community, for example adding a link to the Digital Curation Centre’s (DCC) Disciplinary Metadata guide (<http://www.dcc.ac.uk/resources/metadata-standards>).

Although some of the guidance topics arose from requirements within the University policy they were not the only reason for selecting the areas to be covered. Those who had pioneered support in this area, such as Monash University, Australia (<http://monash.edu/library/researchdata/>), influenced our selection and presentation of the material. Our aim was to provide information that would be of value primarily to new and early career researchers, but would also be of use to Principal Investigators and those more experienced when they reviewed what they needed for their next proposal; that it would draw together information from across the University, thus acting as a hub, but without reproducing what was available elsewhere. We recognised that it would only be one of a number of routes in to support, and needed to complement existing information. We also attempted to keep the material in each section to the equivalent of a double sided A4 page, recognising that there is a balance between information giving and information overload. Topics linked to specific areas of support within the University was written by that area, for example, storage was written by iSolutions, with the DataPool team developing the more generic material. Initial guidance content was circulated for comment to the DataPool Project team and Co-Investigators, members of the University Research and Enterprise Advisory Group (REAG) and Faculty Research Data Contacts. This last group consisted of staff within each of the Faculties who had been given this designation during the development of the University Research Data Management policy.

The Library website was selected as the appropriate location to host this material. It had previously included a small section on data management in its Research Support area. This sat alongside material on Bibliometrics, Open Access, e-Theses. These were all areas that the library supported, not just with information but also with training events and other materials. The offer of research data management support material, guidance, advice and training would be the same, but would draw in a wider network of staff, from different services and areas but including the library, to provide it. It therefore seemed appropriate to develop these more comprehensive research data management pages in this location. Details of the approaches to training for researchers can be found in the *Developing Researcher skills in research data management* report (Byatt et al, 2013).

Use of the guidance appears to have been steady since September (see fig.1 below), when Google Analytics was added to the website as a whole, with the usual vacation dip in December.

**Fig.1 Research Data Management Web site Total Unique Page Views (Google analytics)**

Further usage figures broken down by page breakdown are available in the appendix. The site was introduced to Academic Liaison Librarians and to Collaboration, Bid and Research Support Managers (RIS) in briefings given in the summer so that it would be included in induction events for postgraduate researchers at the start of the academic year and other related training, mentioned in the DataPool blog on Twitter and at research data management events outside the University.

As mention previously, there have been minor changes during that time and in March 2013 more detailed pages were released on “Describing your data effective re-use”.

## 4. Summary and Further Developments

The introduction of the ResearchData@soton service has been a valuable addition in providing support for the researcher in the area of research data management. It has also been very useful in being able to facilitate engagement with different areas within the Institution that support research, this in turn led to joint training events (Byatt et al, 2013). Currently the number of staff available to support this service is limited. The training of additional Librarians to assist has already started and will continue, helping to increase the breadth and depth of subject specific knowledge that can be drawn on.

Requests for help and advice have come from eleven different disciplines and have included History, Modern Languages, Psychology, Health Science, Biological Sciences, English and Medicine with Medicine the most frequent. However even within medicine there has been significant variety covering pure science through to direct patient research, some where all the research is to be carried out within the University of Southampton while others have been joint University/NHS proposals. This has led to a great variety in the detail required within the data management plan.

The majority of RCUK funders have been represented in the bids, as have some charitable organisations and more recently European Union proposals. It has been useful in developing our skills and expertise to have to investigate each of these funders, but more importantly it has drawn together more staff from different areas, thus enabling improved liaison within the University and better data management planning.

We are aware that there are already groups of researchers in the University, in areas such as social statistics, who are quite familiar with the concept of data management planning and have a wealth of experience to draw on. Similarly those researchers looking for NERC funding have established routes of support. Although we have engaged with these groups, for example through DataPool Migration Case study (Wisniowski et al, 2013), they have not yet used this service and perhaps may not have need of it. We are working to incorporate their experience into our support and guidance. We will look to do this through discussion at Faculty or Discipline based events on research data management. Those who have contacted us appear to be less familiar with the issues involved in this area. Data Management terminology appeared to be unhelpful, but difficult to avoid. Metadata seems to be the one that has the most negative effect when first encountered. Although existing guidance is valued, having someone interpret it by giving examples and relating it to the research proposed has been the aspect found to be most helpful and most appreciated.

All areas of data management have been covered. There have been requests for advice on handling sensitive data, keeping data safe when working away from campus, and storage and back-up. The majority of requests have involved iSolutions and the Library in the responses and stronger links have developed, particularly with the Medicine and Health Sciences Business Relationship Manager. This has resulted in opportunities for on-going training within these areas (Byatt et al, 2013). Staff training requirements have also been identified and this has fed in to the training strand of the DataPool project (see Byatt, 2013). Future developments will include building on this model with training for all Business Relationship Manager (already agreed) and with the Collaboration, Bid and Research Support Managers in RIS.

The guidance on the web site has been developed to provide a point of reference for researchers and professional service staff on aspects of research data management, make links to related resources and direct enquiries to the ResearchData@soton service. The various pages all show steady use. Although the guidance covers many of the essential area in research data management there will be a need to add further topics. Work will also be needed to embed links to the information in official University documents, such as handbooks and course materials, as well as in materials for research data management training events. It will continue to be important to maintain our iterative and collaborative approach, developing the content with academic and professional services colleagues as well as improving the links to related resource elsewhere.

In conclusion, the DataPool Project has enabled the development of the guidance and ResearchData@soton desk side service discussed in this report. The service has been well received by those who have used it. Development of the service will continue through the training of additional library staff able to respond to requests. The network of individuals that developed out of responding to requests has led to a number of opportunities to promote research data management and raise awareness of the University policy within the institution amongst both professional services staff and researchers. These have continued to arise with two further requests to meet with staff already planned for May. Support of research data management will continue through the work of the University Research and Enterprise Advisory Group and through the stronger links developed between services and researchers as part of the project.

## References

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>

Byatt, Dorothy (2013) *Supporting Researchers with their Research Data Management: Professional Service training requirements* Southampton, GB: University of Southampton [in Press]

Digital Curation Centre (2013) *Overview of funders' data policies* <http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies> [accessed 18/04/13]

Medical Research Council (2012) *MRC guidance on data management plans* [internet] <http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/datasharing/DMPs/index.htm>

Medical Research Council (2011) *MRC policy on research data-sharing* [internet] <http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/datasharing/Policy/index.htm>

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>

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

Wisniowski, Arkadiusz, Chivers, Martin and Whitton, Michael (2013) *Integrated Modelling of European Migration Database Case Study* Southampton, GB: University of Southampton <http://eprints.soton.ac.uk/id/eprint/350672>

## Acknowledgements

The guidance pages could not have been developed without the assistance of staff from iSolutions, Research and Innovation Services, the Library and other Professional Services within the University.

Our acknowledgement of thanks also goes to the DataPool Co-Investigators, to the Faculty Research Contacts and to the members of the Research and Enterprise Advisory Group who reviewed guidance and offered expertise.

Our thanks also go to those who were contacted via ResearchData@soton and whose help was invaluable in responding to the researcher’s requests.

## Copyright

© 2013 University of Southampton



This work is licensed under the Creative Commons Attribution 2.0 UK: England & Wales License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/2.0/uk/>.

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 and White, Wendy (2013) *Research data management planning, guidance and support: a DataPool Project report.* Southampton, GB, University of Southampton, 16pp

# Appendix

## Research Data Management Web Page Analytics