JISC Final Report

Data Management Planning for ESRC Research Data-rich Investments project (DMP-ESRC)

Project Information			
Project Identifier	*To be completed by JISC*		
Project Title	Data Management Planning for ESRC Research Data-rich Investments project (DMP-ESRC)		
Project Hashtag	@UKDA_RDMSS		
Start Date	1/1/2010	**End Date**	10/5/2011
Lead Institution	University of Essex, UK Data Archive		
Project Director	Louise Corti		
Project Manager	Veerle Van den Eynden		
Contact email	veerle@essex.ac.uk		
Partner Institutions	ESRC		
Project Web URL	http://www.data-archive.ac.uk/create-manage/projects/JISC-DMP		
Programme Name	*Managing Research Data*		
Programme Manager	Simon Hodson		

Document Information			
Author(s)	Veerle Van den Eynden, Laurence Horton, Louise Corti		
Project Role(s)	Project manager, project officer, project director		
Date	April 2011	**Filename**	Finalreport_DMP-ESRC_date2011
URL	*http://www.data-archive.ac.uk/create-manage/projects/jisc-dmp?index=1*		
Access	This report is for general dissemination		

Document History		
Version	**Date**	**Comments**
00.01	12/4/2011	Draft by LH and VVdE
00.02	9/5/11	Edited by LC
00.03	10/5/11	Edited by LB

Table of Contents

1 Acknowledgements

This project was funded by JISC as part of its Managing Research Data programme. It was coordinated by the Research Data Management Support Service at the UK Data Archive in collaboration with the Economic and Social Research Council (ESRC) Policy and Resources Directorate. We particularly thank the three participating centres - Centre for Research on Socio-Cultural Change (CRESC), Centre for Social and Economic Research on Innovation in Genomics (innogen), and Third Sector Research Centre (TSRC), and the New Dynamics of Ageing programme for the cooperation of directors, managers and researchers and their willingness to share information on how research data are managed.

2 Project Summary

This project, funded by JISC under the Managing Research Data Programme, has been strengthening data management and sharing capability in the social sciences community. The UK Data Archive worked with four Economic and Social Research Council (ESRC) large research investments:

New Dynamics of Ageing (NDA) Programme

Centre for Research on Socio-Cultural Change (CRESC)

Centre for Social and Economic Research on Innovation in Genomics (Innogen)

Third Sector Research Centre (TSRC)

We developed data management planning guidance for these research hubs, in line with the new ESRC Research data policy, to ensure that the high quality research data they produce can be made available for wider reuse.

After evaluating existing data management practices by researchers, we helped develop for each centre and programme a tailored set of data management recommendations and a planning strategy. These support existing good practice and direct hubs to resources or tools that address data management weaknesses. By comparing the specific strategies and distilling common elements, we were also able to develop generic data management recommendations for research centres and programmes.

We gave customised training – in response to requests from the hubs – on relevant data management topics: how to prepare research data for archiving; Q&A session for centre researchers on the data management planning recommendations made; recommendations for managing qualitative research data using NVivo software.

Although we found good data management practices to exist, these are idiosyncratic and variable between researchers and largely the responsibility of individual researchers. Many researchers are capable data managers, not necessarily referring to the research or project management practices they implement as 'data management'. However, good practices need to be exchanged much more between colleague researchers and implemented more systematically. Whilst many researchers have a positive attitude to sharing data formally or informally, good intentions need to be translated into more effective data sharing strategies.

Research project leaders and centre directors need to take responsibility for co-ordinating core data management aspects (data security, backup strategy, formatting), a system to keep track of all research data being created or acquired, ensuring data management roles and responsibilities are allocated and providing a framework of standard practices, policies and tools. We delivered various resources for use to help implement such framework. Projects can retain flexibility where needed (e.g. consent practices across disciplines).

Core project outputs are:

- Report on Data Management Practices in the Social Sciences, based on interviews with directors and researchers of past and present centres and programmes
- ESRC approved Data Management Recommendations for Research Centres and Programmes, aligned with the ESRC Research data policy
- Data inventory template and model data management resources library
- Guidance for ESRC applicants and reviewers on data management planning to support the ESRC Research Data Policy, incorporated in online Je-S form and on ESDS website
- Activity-based data management costing tool to help calculate costs of data management above standard research procedures
- New edition (3rd) of the UK Data Archive's Managing and Sharing Data guide, fully revised and with new content on data management planning, strategies for large-scale research investments, and data management for collaborative research
- New content and resources for the UK Data Archive's Create and Manage web pages.

3 Main Body of Report

3.1 Project Outputs and Outcomes

Output / Outcome Type	**Brief Description and URLs (where applicable)** **All project outputs can be accessed via the DMP-ESRC project outputs page**
Report	UK Data Archive, 2010. Data Management Practices in the Social Sciences. Report based on evidence gathered from researchers funded by the ESRC in its various past and present investments. The diversity of data management approaches within the evaluated investments provided important guidance for the development of future data management planning strategies.
Questionnaire	Data audit questionnaire used to record researchers' data resources and their management. Available on the JISC RDM wiki.
Minutes	Minutes and recordings of all project meetings with centres and programmes; as well as interviews with centre directors and researchers to assess data management practices. Internal.
Internal reports	Assessment of data management practices at TSRC, CRESC, Innogen and NDA. Internal to each centre / programme
Internal reports	Data management strategy for TSRC, CRESC, Innogen and NDA. Internal to each centre / programme.
Webpage	Project web page containing information on the project aims and structure as well as outputs
Training	New Dynamics of Ageing, TACT3 project training event: Data Archiving Explained. Brunel University, 27 Sept 2010.
Training	Third Sector Research Centre, Data management for qualitative longitudinal research using NVivo. 15 Feb 2011.
Training	Innogen data management planning recommendations and Q&A session for researchers. Innogen Research Programme Retreat, York, 16 May 2011.
Tool	Activity-based data management costing tool to help calculate costs of data management above standard research procedures.
Tool	Data inventory template
Tool	Model data management resources library and exemplar protocols and documents list
Web info	Variety of new online data management guidance on the UK Data Archive's Create and Manage web pages: data management planning, data management strategies for centres and programmes, data management for collaborative research.
Report	ESRC approved Data Management Recommendations for Research Centres and Programmes, aligned with the ESRC Research data policy. Includes guidance on areas such as consent for reuse, data ownership, anonymisation, documenting data, formatting and file formats, version control, safe storage, transmission, and destruction of data; and suggestions for centre / programme providing data management planning framework by allocating roles and responsibilities, maintaining a centre-wide data inventory and developing a data management resources library.
Seminar	Data Management Planning and Practices for Research Centres and Programmes, Royal Statistical Society, 4 May 2011. This seminar brought together researchers, directors, research managers and support staff from ESRC-funded Research Centres and Programmes with funding councils and data services, to share knowledge and good practice and discuss the roles and responsibilities of the various stakeholders in data

	management - directors, researchers, support staff, institutions, funders and data services to implement effective data management and sharing. The UK Data Archive outlined recommendations from the project.
Videos	Short videos of staff from participating research hubs, sharing key messages at our final seminar on Data Management Planning and Practices for Research Centres and Programmes, Royal Statistical Society, 4 May 2011.
Guide	UK Data Archive 2011. Managing and Sharing Data – best practice for researchers. Third edition. This 3rd edition includes new content on data management planning, data management strategies for large-scale research and data management for collaborative research.
Policy-related guidance	Guidance for ESRC applicants and reviewers on data management planning to support the ESRC Research Data Policy, incorporated in online Je-S form and on ESDS website. Also general data management and sharing guidance is being prepared for the ESRC website, to accompany the Research Data Policy.
Dissemination	Various dissemination activities to share the findings and lessons learnt. See annex.

3.2 How did you go about achieving your outputs / outcomes?

3.2.1 Aims

The Data Management Planning for ESRC Research Data-Rich Investments project (DMP-ESRC) aimed to work with selected ESRC centres and programmes to:

- evaluate existing data management practices amongst researchers in the social sciences community, in particular in large research investments, and identify needs
- help develop and integrate effective data management planning procedures and tools into the research lifecycle
- increase the data management and sharing capability of research hubs within the social sciences through support and training
- develop a data management cost model for researchers in the social sciences
- disseminate findings and lessons learnt to the academic community.

The project took place within the context of recent changes in the ESRC Research data policy This policy requires that research data should be openly available to the maximum extent possible, through long-term preservation and high quality data management. The 2010 policy introduces the concept of data management and sharing planning as part of all research grant applications, to enable better quality data ready for depositing with ESDS and sharing. A plan addresses consent and plans for data reuse, anonymisation and data security, data management roles and responsibilities, and strategies for addressing data management difficulties.

3.2.2 Methods

The project approached 7 ESRC centres and 3 programmes for participation early in 2010. Seven expressed an interest in cooperating and we selected 3 centres and one programme on the basis of challenging data management environment, stage of the centre/programme lifetime, research domains and potential complexity of data sharing.

- Centre for Social and Economic Research on Innovation in Genomics (Innogen)
- Centre for Research on Socio-Cultural Change (CRESC)
- Third Sector Research Centre (TSRC)
- New Dynamics of Ageing (NDA) Programme

The ESRC Centre for Social and Economic Research on Innovation in Genomics (Innogen) runs from 2002 to 2012. It is part of the ESRC Genomics Network and studies the evolution of genomics and life sciences and their social and economic implications and is based at the University of Edinburgh and the Open University. Law, economics and social sciences researchers engage in research projects in the UK, Africa, China and India.

The Centre for Research on Socio-Cultural Change (CRESC), 2004-2014, is based at the University of Manchester and the Open University. Its mission is to analyse socio-cultural change in the context of socio-technical innovation, economic insecurity, and cultural diversity, with the intention to recognise different definitions and approaches to culture in its interface with processes of social change. CRESC research covers quantitative reuse of secondary data (e.g. longitudinal survey analysis) and qualitative research (ethnography, interviewing, audio and visual data).

The Third Sector Research Centre (TSRC), 2008 - 2013 is a collaboration across the Universities of Birmingham, Southampton, Kent and Middlesex. TSRC will bring together experts from a range of disciplines to develop a research programme that will lead to improved understanding of the key patterns, processes, and impacts of developments in the sector. This will strengthen the evidence base for policy towards the sector. TSRC will collaborate with, and offer a wide variety of services to, those working in and supporting the voluntary sector. In addition, TSRC co-ordinates the work of three Capacity Building Clusters which will support and enhance research capacity within the sector.

The New Dynamics of Ageing (NDA) programme is a cross-council collaboration seeking to improve the quality of life of older people. It runs from 2005 until 2012 and is funded by the Economic and Social Research Council (ESRC), Engineering and Physical Sciences Research Council (EPSRC), Biotechnology and Biological Sciences Research Council (BBSRC), Medical Research Council (MRC) and the Arts and Humanities Research Council (AHRC). The emphasis is strongly on multidisciplinary and cross-disciplinary research. Its funding is £22 million, with 43 percent contributed by the ESRC. The programme currently consists of 35 projects across 62 UK higher education institutes. NDA projects cover 47 disciplines, the main ones being psychology, sociology, health sciences and primary care.

After introductory meetings were held with directors and core staff of each participating hub, we interviewed researchers, directors, centre managers and support staff from 26 projects across collaborating ESRC centres and programme between April and September 2010, gathering information on how research data are created, stored, handled, managed and shared. Researchers were selected on the basis of presenting challenging issues of data management, methodological diversity, cross-institutional (and international) collaboration, varying stages of research completion and representation of a centre's research diversity. Additional one-off interviews were held with the director and/or staff members in the Centre for Competition Policy at the University of East Anglia, the Centre for Economic and Social Aspects of Genomics at Cardiff University, and the Centre on Migration, Policy and Society at the University of Oxford, to capture information on data management practices in completed initiatives.

At TSRC, researchers representing ten of the centre's core research programmes at the Universities of Birmingham, Southampton and Middlesex were interviewed. For CRESC six researchers from the University of Manchester and the Open University were interviewed. Six Innogen researchers at the University of Edinburgh or Open University were also interviewed. Researchers from four projects in the New Dynamics of Ageing programme with a significant social science focus were interviewed: ProjectMIMA, TACT3, HALcyon, and SomnIA. Interviews were conducted with principal investigators on these projects and other members of the project were also often present and contributed to these interviews, or were available to answer questions.

All interviews were undertaken face-to-face and followed a data inventory questionnaire designed by the DMP-ESRC project, based on and adapted from the Digital Curation Centre's Data Asset Framework, the Rural Economy and Land Use Programme Data Support Service Data Management Plan and the Australian National University Information Literacy Programme's Data Management Plan. This inventory questionnaire also acted as an interview schedule allowing scope for discussion and supplementary questions where appropriate. Topics covered included ethics and consent for data sharing and re-use, copyright, data anonymisation, data documentation, data re-use and archiving, data formatting, file formats, data validation, version control, data security/storage/back-up, data transmission and destruction of data and documentation.

At the same time, interviewees were asked to provide information or estimates on the cost of various data management activities, considering the cost of data creation (data collection and data entry/transcription) and the cost of preparing data for sharing and re-use. Data management costing information was also gathered from some Timescapes researchers, by asking for time estimates to prepare qualitative longitudinal data for archiving, and from the Relu Data Support Service (Relu-DSS). When data from Relu projects are deposited for archiving at the UK Data Archive, they are prepared for sharing by the Relu data processor. Varying levels of preparation may be needed before data are ready to be archived, depending on the amount of preparation that has been done by researchers. For some data collections, Relu-DSS carried out full anonymisation of transcripts and audio recordings, for other projects simply checking of anonymisation carried out by researcher. For quantitative data, quality checks and data labelling were needed. The Relu data processor kept track of the amount of person hours spent on preparing data for archiving, which could then be compared to total project costs.

Interviews were recorded using a digital voice recorder for note taking purposes and interview notes written up immediately following the interview. Projects also occasionally provided documentation from the project as examples to illustrate data management practices.

3.2.3 Data management planning strategies

All information gathered was organised into a spreadsheet for each hub. This served to first analyse and synthesise the existing data management practices and then to develop specific data management planning recommendations for each centre / programme. The centre / programme gave feedback on the suitability and feasibility of the recommendations, so we could adapt them where needed. The recommendations were also shared with the respective ESRC case officers and the ESRC Principal Policy Manager (Research Resources) Policy and Resources Directorate.

Based on the recommendations developed for the three participating centres and programme, generic data management recommendations for research centres / programmes were developed, in line with the new ESRC Research Data Policy. This formed the basis for developing specific guidance for ESRC grant applicants on data management planning, to support the ESRC Research Data Policy on the ESRC website. The findings were also used to develop best practise guidance for research centres and programmes on data management planning and procedures, which has been included in the Third Edition of the UK Data Archive's guide Managing and Sharing Data – best practice for researchers. This guidance has also been added to the Archive's online data management guidance.

3.2.4 Data management training

Each centre / programme could ask for any training provision relevant for their purposes and needs. As part of the data management recommendations made to a centre or programme, suggestions were made for data management training. But it was left to each centre / programme to request or suggest training to suit their needs.

TSRC were keen to receive training on how data management for longitudinal qualitative research could be applied in NVivo software used to organise, code and analyse data. This required us to liaise with specialists in NVivo and seek solutions to practically apply data management for that particular software package. Our suggestions were discussed against the practices and needs of researchers during interactive session with TSRC researchers.

NDA sought detailed information on the practicalities of archiving research data at the UK Data Archive and how to prepare data for deposit.

Innogen requested us attending their 2011 research programme retreat to discuss the recommendations directly with all researchers and be available to answer questions researchers may have.

3.2.5 Costing tool

As preparation for developing a data management costing model, we attended the JISC MRD Costs, Benefits and Sustainability workshop and gathered ideas on costing models during the MRD Programme Progress workshop.

Information on the cost and time required for various data management and sharing activities was gathered during data management interviews with researchers of 30 projects in participating centres and programmes. The researchers were asked to estimate the time or cost needed for activities related to: data collection, data entry and transcription, data validation and documentation and the cost of preparing data for archiving and re-use.

We then discussed within the project team the purpose of costing data management as part of the research process. Data management can be seen as an integral part of the entire research process, thereby meaning that all data-related research activities can be seen as data management and therefore be costed. On the other side, a more useful data management cost to know is the additional cost of making research data shareable beyond the primary research. Since all research data resulting from ESRC-funded research are typically archived at the UK Data Archive, the data management costs only need to take data management activities until the point of depositing data with the Archive into account.

This latter costing approach was used as the basis for developing the tool, considering that the main purpose of costing data management is to:

- help researchers understand what to consider in research to make data shareable so researchers request adequate funding for creating shareable data
- provide funding councils with better estimates for the data management costs of enabling wider access to data
- help data service providers reduce the costs of data processing for archiving and re-use, through properly costed data management plans.

3.3 What did you learn?

3.3.1 Data management practices in research hubs

From the interviews it was apparent that there exist many good data management practices within the participating research centres and programme. Researchers are capable data managers, although they may not refer to the research and project management practices they implement as ‘data management’ or ‘DM planning’.

However, good data management practices are idiosyncratic and variable between researchers and not often shared between colleagues within the same centre.

Researchers generally have an open attitude to data sharing and reuse, either during or at the end of a project – and most have explicit commitments to sharing data, either formally or informally. For some, obstacles to data sharing are a perceived lack of interest in researchers re-using data. Other challenges to data sharing are posed by cross-disciplinary projects and cross-agency funded projects, where it is not always clear whether the ESRC data policy or another data sharing policy applies or not.

Despite such general positive attitude towards data sharing amongst researchers, there is no coordinated or systematic data sharing or archiving strategy. Reviewing data deposit rates from ESDS on the amount of datasets archived from particular grants, research centre grants tend to score poorly, with relatively few datasets archived for the amount of money invested in research. This may well be because the mandate to offer data for archiving only kicks in when the grant ends, which for centres is at least 5 and often 10 years after the grant started. After such long time it is very likely that poorly managed data are very difficult to find and prepare for archiving, many researchers have moved on, etc. In comparison, smaller and shorter grant have better data deposit rates.

A centre or programme director may coordinate data management as part of research management, but this is not usually the case. Only TSRC had a data management committee in place, consisting of senior management and selected representative researchers, meeting six-monthly to discuss pertinent issues. Other pressures take precedence for centre management, such as outputs and impact through high quality research publications and communications. In practice most aspects of data management are the responsibility of individual researchers and the practices vary enormously

between researchers. Although data management is not usually formally planned, certain aspects form part of research, e.g. ethical review procedures, data collection strategies or data preparation activities. Where cross-site or cross-national data collection activities exist then data management is (and has to be) more tightly coordinated.

The centres and programmes could see clear benefits from the start of the project of the centre providing a data management planning framework of information, best practices and policies to its researchers and welcomed the opportunity to participate in the project. These benefits of a planning framework - and awareness of the disadvantages of not having one - thus became the principal recommendation for data management planning. Since centres and programmes benefit from dedicated central research coordination, this can be used to provide a framework in planning and implementing data management and sharing activities.

3.3.2 Gaps in UK Data Archive, ESDS and ESRC guidance

By analysing the data management needs of researchers for their research practices, the project identified gaps in the existing UK Data Archive's data management guidance, in particular on:

- Data management planning
- Data management roles and responsibilities
- Costing data management
- Guidance for (external) transcribers
- Folder structures
- Data transmission and encryption
- Collaborative research environments and file sharing

Guidance was therefore developed on these topics and added to the data management pages on the Archive website; as well as included in the Managing and Sharing Data guide, edition 3. This evidence-based guidance arose directly from synthesising feedback from the work with sites. This project certainly deepened our understanding of research practices in centres and means that our advice is now more grounded in real research practices and can better address the needs of researchers.

At the same time as this project was proceeding, we were working closely with ESRC to help formulate the new Research Data Policy and help draft guidance for research applicants and data management plan reviewers. The timing of these activities meant that guidance arising from the project site work fed directly into ESRC recommendations in a timely way. Our expanded advice for ESRC applicants is now much improved, especially where we can now refer to real case studies. This could only have happened with ESRC being directly involved in the project, and we commend the JISC MRD Programme Director for having the foresight to enable this opportunity.

3.3.3 Costing data management

It was found that overall data management and data archiving is not costed in or planned much during a centre's planning stage, unless the centre has a specific remit to deliver data for reuse, like the Institute for Social and Economic Research (ISER) or the Centre for Longitudinal Studies (CSL).

At the same time, researchers are keen to know how to cost data management when preparing grant applications and ESDS receives many queries on this topic.

Overall researchers interviewed found it hard to cost data management activities, as many activities are an integral part of standard research activities and data analysis. Identified high cost activities are transcribing and anonymising qualitative data and cleaning and verifying quantitative data.

The information from the researchers was combined with the Archive's expertise in the measures needed to make research data shareable and re-usable.

Key is that when researchers leave preparing data for sharing until the end of a project, the cost is far too high due to competition with publishing and seeking future project funding. Data management should therefore be carried out earlier on in a project.

An activity-based data management costing tool was therefore developed, to help social science researchers measure the additional costs – above standard research procedures and practices – that are needed to make research data shareable beyond the primary research team. The tool is based on

identifying all applicable data management activities and steps required to make data shareable, based on a data management checklist, then costing each activity in terms of people's time or physical resources needed such as hardware or software.

Various stakeholders, staff and researchers at participating centres and programmes, and participants of the project's seminar on Data Management Planning and Practices for Research Centres and Programmes have been asked to review the costing tool and provide feedback. The tool will be further developed / adjusted based on this feedback.

3.3.4 Recommendations for centres and programmes

Based on the four data management planning strategies developed for each participating centre / programme, and after feedback from each and further discussions with relevant staff, the following are an overview of key data management recommendations for research centres and programmes in the social sciences. Despite the wide disciplinary and methodological diversity of research and researchers at the participating hubs, the data management gaps identified and recommendations made for each centre / programme were remarkably similar in many aspects.

Topic	Recommendation
DATA MANAGEMENT STRATEGY *Choosing the level of centre- or programme-level coordination and decision-making in planning and implementing management and sharing of research data*	Level of centralisation or researcher autonomy for data management determined by type of research activities, methodological and discipline diversity, size of research hub, single or cross-institutional entity Allocate clear roles and responsibilities for all data management aspects; designate data management coordinator for each project Develop a data management resources library Create a data inventory Recommended file structures for projects
ETHICS AND CONSENT *Informed consent for people to participate in research and for use of information collected*	Consent not prohibiting data sharing Gain consent for sharing, preservation and re-use of research data
COPYRIGHT *Intellectual property rights of data*	Identify source and ownership of third party data Establish conditions of use, copyright constraints and redistribution rights Statement on intellectual property rights of generated data
ANONYMISATION *Procedures to protect the identify of individuals, organisations or businesses and avoid disclosure*	Plan and anonymise data early in research Create anonymisation log of edits, replacements, removals or aggregations made
DOCUMENTATION *Material explaining how data are created, what they mean, their content and structure, and alterations done*	Create and sustain comprehensive documentation of data creation methodology Define specialist terms and acronyms used
RE-USE	

Plans for sharing and future use of data	Know all relevant funding council data sharing policies Know data centres and I institutional repositories to preserve your data and distribute them for re-use
DATA FORMATTING *Layout, editing and structuring of data*	Guidance for researchers and transcribers on transcription procedures and use of uniform editing and layout Use coversheet for transcripts of qualitative interviews Use non-disclosure agreement for administrative staff and transcribers accessing confidential data List of recommended transcribing services
FILE FORMATS *Formats and software in which research data are created, analysed, stored or preserved*	Encourage use of standard or open lossless formats
VERSION CONTROL *Checks and procedures to establish which version of a file is the most current*	Establish procedures to control which version of a file is the agreed shareable version, or master version
STORAGE AND BACK-UP *Steps to store data and back-up procedures to make copies to restore originals in case of data loss*	Know your institutional arrangements and capacity for data storage Know your institutional procedure and regularity of data back-up, especially for remote and cross-institutional working
FILE SHARING AND STORAGE *Strategy for sharing and storing research data and providing shared access*	For cross-institutional collaborative research, assess the feasibility of adopting a virtual research environment or file sharing and storage system
TRANSMISSION *Methods of moving data files from one location to another*	Only send anonymised or non-sensitive data in unencrypted form Encrypt confidential or personal data before transmission or password protected files Explore secure storage and sharing files and folders with others across the Internet using file synchronization, virtual research environments or file sharing software
DESTRUCTION *Procedures for safe and secure erasing of data at the conclusion of research project or when needed*	Policy on retention and disposal of data

3.3.5 Pertinent data management needs

Two important needs of data centres and individual researchers emerged.

For collaborative, cross-institutional research projects there is a clear need for user-friendly collaborative research environments that require little technical knowledge from the part of the researchers, yet are trustworthy and secure for data handling. Systems should enable shared access to data files (including external access) and file sharing, whilst ensuring data security and version control. Whilst one centre was very pleased with their use of MS Groove 2007, none of the other centres or projects had been able to find a satisfactory system for their needs.

At the centre- or programme-level there seems a clear role for the research hub to provide a coordinated data management of shared best practices, guidance, templates and resources and promote the sharing of good data management practices that exist amongst their researchers. This framework should ideally be initiated and supported by the director, who has ultimate contractual responsibility, and managed by the centre manager or someone with a central administrative responsibility. Populating the data management resources library should be a first key activity, followed by nominating a data manager for every research project or sub-group and ensuring that they meet to discuss common topics – either as a separate data managers group or a standing item on regular centre researchers meeting agendas. This data manager role is likely to be taken on by a key research officer looking after or handling after data.

3.4 Immediate Impact

3.4.1 Impact for participating research centres

As a result of the specific data management planning recommendations made, both the TSRC and Innogen already implemented new data management measures and shared those with participants at the Data Management Planning Strategies and Practices seminar.

TSRC already had a data management committee in place before their engagement in this project. As a result of their engagement in the project and the recommendations made, various new measures were taken:

- Data management is a rolling agenda topic at all team meetings.
- Sharepoint had already been chosen as the centre's working platform across the four centres. Creating a dedicated space to share work and documents has made data management much easier. TSRC developed a shared data management space so researchers and others can find guidance or templates at any time or contribute to the space. The data management space is divided into separate folders: Research Areas, Ethical Review, Transcription, Data Management and References.
- In each research stream, a lead person is nominated to report to the centre's Data Management Committee. A folder is created for each research stream within the Research Areas folder. Researchers can create shared working documents in this area or submit completed work. Version control can be executed through this means.
- The Ethical Review folder contains a library of templates and forms on consent, confidentiality, ethics compliance, etc. as well as minutes and correspondence for the Research Ethics Committee.
- In the Transcription folder guidance documents can be found on coding and numbering systems to be used by transcribers, as well as details of the centre's transcription service with evidence of compliance from all transcribers. For each research stream and projects within streams, project logs are held as well as transcriptions, recordings, evidence of ethical approval and project-related documents.
- The References folder holds policies of procedures for the Universities of Birmingham and Southampton, the ESRC and the UK Data Archive.
- Work is in progress to further develop the SharePoint areas and write recommended policies and procedures
- Researchers are required to report to ESRC via SharePoint to encourage engagement with data management and SharePoint.

Innogen is proceeding with developing a data inventory for the centre and are committed to start adopting the recommendations made.

CRESC had a change of director during 2010, which disrupted the engaged dialogue over the recommendations and strategy. Whilst the initial centre meeting was very insightful to the challenges of data management in a large and diverse research centre and researchers were very cooperative during interviews, the change of leadership meant that the centre lacked a senior advocate to progress the dialogue around data management planning recommendations.

3.4.2 Impact for ESRC

We have provided significant and timely input for the development of the new ESRC Research data policy, which includes a requirement for data management plans, and the format and focus of such plans, ensuring that policy and data management planning requirements are informed by researchers' practices. We have provided to ESRC guidance for researchers completing a data management plan and for peer reviewers evaluating a plan. Based on the findings and recommendations of the projects, we have also drafted data management guidance specifically for the ESRC website, to support the policy.

We think that without this project, and the direct involvement of ESRC in it, ESRC may not have been able to move so quickly and efficiently on this Policy work – which has been unbelievably fast in the timeline of Research Council activities.

3.4.3 Impact for the UK Data Archive

At the UK Data Archive the immediate impacts of the project have been numerous and productive:

- Enhanced and expanded online data management guidance with new dedicated sections on data management planning and strategies for large investments, as well as additional guidance on costing data management, folder structures, data transmission and encryption and collaborative research environments and file sharing.
- More detailed data management planning guidance and directly supporting tools, e.g. a ready-to-use Data Management Resources Library and exemplars/documents, and actual case studies to cite and use when advising ESDS researchers
- Data management costing tool and guidance, for which there was an urgent need from grant applicants
- New edition of the Managing and Sharing Data guide: 1000 copies of the 1st edition (2009) and 2000 of the 2nd edition (2009) of this popular guide have now been distributed within the research community; 3000 copies have been printed for the 3rd edition; in addition the online PDF version of the guide has been downloaded 525 times since June 2010 (when web traffic tracking on the current Archive site started). Since the publication of the 3rd edition of the guide on 3 May 2011, almost 500 copies have been distributed within 10 days via requests, with bulk requests from the Third Sector Research Centre, DATUM for Health, Newcastle University Library, the Institute for Social and Economic Research and the Archaeology Data Service; distribution at data management events and the online version was downloaded 100 times in 8 days.
- Networking with other projects of the MRD programme: this project enabled us to meet great people with complementary skills. Collaboration has been set up with the FISHnet project, with whom we would like to explore closer cooperation, based on their technical solution for data sharing.
- The project has resulted in close collaboration with some of the JISC Research Data Management Training projects (RDMTrain). Three of these have sought our involvement – inviting us to speak at seminars on our ethical expertise and to provide advice, especially for qualitative data management. The RDMTrain projects have also benefited from the Archive's concurrent ESRC Research Development Initiative (RDI) project on Data Management Training for Researchers, aimed at capacity building for all levels of research community. Our work on delivering over 15 data management training sessions in the past year has enabled us to support RDMTrain projects with quite specific grounded advice. The DataTrain and DATUM for Health projects visited us for advice and used the Data Management Planning checklist we had prepared for use by PhD candidates, after finding the DCC checklist too complex for junior trainee researchers. We were commissioned by EDINA to deliver a training module on qualitative data preparation for NVivo for the Research Data MANTRA. At the same time the DMP-ESRC project was writing data management guidance for qualitative data using NVivo for TSRC. Synergy between those two projects has thus enabled us to deliver evidence-based advice in a much broader arena.
- A collaboration on data management has been established with the Digital Curation Centre (DCC) with the Archive contributing towards some of the DCC's data management planning roadshows.

- Experience gained in the use of social media (blogging, tweeting) for project communication, which the Archive had previously not been involved with.

Overall, many beneficial experiences have been shared with other JISC MRD projects, both at JISC MRD events and by contributing towards activities and events of other MRD projects (see dissemination activities).

3.5 Future Impact

The greatest impact we believe will be on ESRC award holders – and especially on Research Centres – who, as a result of the new guidance and advice capacity of the ESDS will be better placed to take up workable and efficient data management activities. We would expect this to lead to better and more data sharing. This is being tracked via the data that ESDS records on data deposits for all ESRC awards, which is linked with ESRC's information on all awards generating data.

The impact should spread to any kind of research centre/hub for which our strategic guidance and tools can be taken up, irrespective of discipline. This also applies to centres outside of academia where we have noted interest in our materials and advice arising from this project. For example, a member of the Social Research Association voiced interest on behalf of local councils.

We also believe that we influence data management and sharing at a cross-disciplinary and international level. Our data management and data management planning guidance and resources are also frequently cited by the research data management and sharing community. Analysing traffic to our web pages on managing and sharing research data shows that many users are referred to our web resources from the websites of MIT Libraries, Edinburgh University, Council of European Social Science Data Archives (CESSDA), Manchester Metropolitan University, Purdue University, International Association for Social Science Information Services and Technology (IASSIST), Digital Curation Centre, Wellcome Trust and Columbia University.

For large and long-term investments such as centres and programmes, discussions with ESRC have been held already on how data sharing can be driven more. ESRC are already encouraging all ongoing large investments to also develop a data management plan. Other ideas that have been discussed with ESRC are considering a cycle of data archiving throughout the lifetime of an investment, rather than leaving data archiving until the end of an investment (when researchers are aware of the data archiving mandate as part of final reporting); and case officers following a centre's data management and sharing practices with hubs reporting back to ESRC on data managing and sharing as part of more general reporting obligations. If data management roles and responsibilities can be correctly assigned before any contract is awarded, impact on a centre's data sharing activities is likely to be noticeable. The Archive will continue this discussion with ESRC and the data deposit rate of large investments is being monitored as part of general ESDS monitoring of data deposit rates from all ESRC awards.

There has equally been discussion on the UK Data Archive providing data management training for ESRC Centre case officers to support a strategy of case officers supporting data management activities within investments.

The final project workshop held on 4 May 2011 attracted several strategic agents – some high profile centre directors, centre managers and those looking at data policies for other funders and Research Councils. The Medical Research Council indicated to be very keen to use our advice and models as the basis for their own advice for centres and large projects, which we are delighted about. Thus impact has already spread to other disciplines as a result of our early dissemination activities. Although the UK Data Archive and ESDS have been leading the way in data management planning for Research Councils in the UK for some time now, this project has definitely helped consolidate knowledge and expertise.

We have already embarked upon communications for advertising our new Managing and Sharing Data guide (with results as shown in 3.4.3) and we expect more data support hubs to follow our guidance. We expect many more instances of 'plagiarism', but it is all positive plagiarism that spreads the word, which has obvious impact!

We hope that we can claim impact for JISC and for helping produce timely outputs from this JISC programme.

4 Conclusions

We welcome the fantastic opportunity we had to be able to work closely with the ESRC, with selected ESRC research centres and programmes and with the JISC MRD community - providing much topical synergy - over the last 15 months around data mangement planning strategies for research hubs.

Not only have we been able to collaborate with important stakeholders in this emerging field of research support, but more importantly, this project has provided us with the opportuniy to work very closely and openly with numerous researchers, who are the people handling research data on a daily basis and needing the capacity, resources and tools to manage and share those effectively.

Thanks to this fruitful collaboration, we have been able to gain much expertise and essential knowledge around data practices. This is essential to ensure that the guidance and advice we provide to the research community is grounded in real research practices and the needs of researchers. We believe this is benefitting us, the ESRC and the ESDS, as well as the wider research community in the UK.

This collaboration has been very timely, as the ESRC developed their new Research Data Policy in 2010 and we were able to provide significant input towards the development of this new ESRC Research data policy, in particular the requirement for data management plans and the format and focus of such plans, ensuring that policy and data management planning requirements are in line with researchers' practices. We provided to ESRC guidance for researchers completing a data management plan and for peer reviewers evaluating plans in research applications.

As a direct result of the project we have been able to develop a range of data management planning support resources, both for the participating research centres and programmes, as well as more widely for research hubs and researchers in general. Resources developed include data management planning recommendations for participating research hubs; generic guidance and practical solutions for all large investments; a data management costing tool; and enhanced and expanded online and published data management guidance with new dedicated sections on data management planning, strategies for large research hubs, data transmission, file sharing and Virtual research environments, and costing data management. This guidance is all grounded in real research practices and based on realising the gaps in existing guidance from analysing the data management needs and practices of researchers.

Our overall finding is that whilst good data management and sharing practices existing within the social sciences research community, this is dispersed and idiosyncratic. There is much scope for research hub directors to provide leadership and for good practices to be exchanged between researchers.

Centres can provide a framework of data management planning support resources and tools, coordinating a uniform approach to data management with central policies on various data aspects; establishing who owns data and keep track of them in time, especially when researchers leave the hub; provide central storage for research data; and ensure that all researchers and staff are aware of duties, responsibilities and funder requirements regarding research data, with easy access to relevant information. When researchers share good practice and data management experiences with each others, this can build capacity for the centre

This project was strongly focussed on some of the core activities of the UK Data Archive, the ESDS and its collaboration with ESRC. The findings and expertise gained will therefore contribute significantly to the ongoing activities of the Archive.

Overall we consider the JISC Managing Research Data programme to be excellently managed and we thoroughly enjoyed (and benefitted from) the collaboration and synergy between JISC MRD projects. We closely collaborated with some MRDTrain projects and with the DCC, with plans for ongoing and future collaboration

5 Recommendations

5.1 Practical recommendations for research centres and programmes

At a very practical level, based on the findings of the project we would like to recommend to all research centres and programmes that providing a data management framework of resources, practices and tools has benefits:

- researchers can share good practice and data management experiences with each others, therefore building capacity for the centre
- establish a uniform approach to data management and central policies on various data aspects
- a centre can establish who owns data and keep track of them in time, especially when researchers move
- data can be stored at a central location
- ensuring that all researchers and staff are aware of duties, responsibilities and funder requirements regarding research data, with easy access to relevant information

Research hubs should consider:

- the Director taking a lead on developing local policy for their centre which gives a high level statement to the effect that: “We manage and share data - and we do it well “
- assigning data management responsibilities to named individuals, which at minimum is a centralised administrative coordinator whose job it is to set up a data inventory for the centre, a data management resources library and assign research project-based data contacts in the centre who meet regularly
- make data management and sharing a recurring agenda items for team meetings
- developing standardised forms for the data management resources library, e.g. for consent procedures, ethical review, data management plans
- developing standards, e.g. data quality control standards, data transcription standard, confidentiality agreement for data handlers
- file sharing and storage procedures
- security policy for data storage and transmission
- data retention and destruction policy
- data copyright and ownership statement for the centres and for individual researchers
- standard data format recommendations
- version control and file naming guidelines
- information on funder requirements or policies regarding managing and sharing data that apply to projects or the centre
- research data sharing strategy, e.g. via institutional repository, data centres, website

5.2 Recommendations for ESRC

What emerged is that research centres, which typically have a lifetime of 5-10 years, are only required to archive data at the end of their lifetime. This often means that relatively few data are archived from such large investments, compared to smaller research grants. Centre contracts need to be more closely scrutinised by ESRC from a data managing and sharing perspective. We suggest the centre

Case Officer takes on this role. Centres should be expected to have a short section at the end of their Annual Report on progress with data management planning and data sharing activities.

Centres attract external grants that bring in money from other funders, which does not sit under the ESRC Contract. Even if there is no formal requirement to share data from these other awards, ESRC should advise that as it provides the infrastructure and bulk of money for supporting the Centres as going concerns, all data should be treated equally and fall under centralised data management coordination.

ESRC may want to consider developing a Data Management Charter for directors of large investments, which would outline expectations on the implementation of data management planning, effective data management and the reporting on data management and sharing practices.

5.3 Recommendation for JISC

We commend JISC and particularly the MRD programme manager for running the MRD programme so brilliantly, in a very pro-active and engaging way, creating close and interactive cohesions between all MRD projects and the wider stakeholder community. It has been a pleasure to engage with it.

For collaborative, cross-institutional research projects there is a clear need for user-friendly collaborative research environments that require little technical knowledge, yet are trustworthy and secure for data handling. JISC may want to consider how a good solution could be developed or supported for the UK HEI. During the unconference sessions at the JISC Managing Research Data (International) programme workshop suggestions were made that the SURFconext and colwiz tools are potentially interesting tools, which JISC may want to consider evaluating.

We would like to advocate using the next data management call to work on embedding the workable practices arising out of the MRD programme. For example, close collaborating with researchers on the ground and with a Research Council and its sharing policy delivers very effective results, as there is obvious ‘clout’.

We would also like to recommend that there is a need for data management planning support for non Research Council funded research investments that have no data policy context to drive data management.

We would recommend that JISC consolidate findings from this MRD programme and ensure that duplication of effort for ongoing data management support activities in the UK does not happen. We have been aware of other services providing similar support activities. Coordination of support activities is needed so that researchers know where to obtain the most relevant advice and so that activities are coordinated efficiently.

6 Implications for the future

Already we are exploring ways to share experiences with audiences beyond academia. A recent data management workshop was held jointly with the Social Research Association with various researchers from charities, private companies and councils attending. Through a small knowledge exchange grant from the University of Essex, Archive expertise on data management and preservation is also being taken to local non-academic bodies that collect data. This will draw on tools developed and findings from DMP-ESRC.

The UK Data Archive was part of a recent bid to the JISC University Modernisation Fund call, which was based on collaboration made through the JISC MRD programme. We collaborated with King’s College’s FISHnet project and although the bid was not successful, it demonstrated excellent collaboration. Future work to combine work from the FISHnet prototype and our own data management hub infrastructure tools will be taken forward, if we can attract funding. We feel this is a very powerful synergy not to be wasted, and has the capacity to feature as an easy-to-use, low barrier tool in every research centre.

One avenue for sustainability is to explore how data management training can be coordinated within general user training and support for commonly used data analysis software packages (similar to the

NVivo guidance we developed). By focussing on ensuring that researchers can implement good data management practices within the analyses software used, data management can be embedded early on in the research cycle.

We also want to consider with ESRC how to capture the longer term impact of data management planning. Detailed costing of the impact of the introduction of data management plans is likely not to be easy or possible. A better approach will be before and after comparisons, testimonies from PIs, even assessments at the UK Data Archive whether data are deposited in a better shape for sharing.

Finally, we would very much like to help embed the good practices in all ESRC Centres. This mission probably needs additional dedicated funding as ESDS does not have staffing to support each centre. We are losing our JISC project officer, who is moving to a German data archive where he will take his experiences gained for European projects. We will continue to explore available funding that will allow us to take on a proactive role in embedding and promoting good data management practices in research, which we believe to be very effective.

7 References

Australian National University (2010) Information Literacy Programme's Data Management Plan, available at http://ilp.anu.edu.au/dm/dmp_template.pdf accessed 24 March, 2011

Cabinet Office (2008) Data Handling Procedures in Government: Final Report, available at http://www.cesg.gov.uk/products_services/iatp/documents/data_handling_review.pdf, accessed 08 March, 2011

Charlesworth, A. and C. Rusbridge (2010) Freedom of Information and Research Data: Questions and answers, available at http://www.jisc.ac.uk/publications/programmerelated/2010/foiresearchdata.aspx accessed 10 January, 2011

Digital Curation Centre (2010) Funders' Data Polices, available at http://www.dcc.ac.uk/resources/policy-and-legal/funders-data-policies accessed 10 January, 2011

Digital Curation Centre ,(2010) Data Asset Framework, available at http://www.dcc.ac.uk/resources/tools-and-applications/data-asset-framework accessed 24 March, 2011

Economic and Social Research Council (2000) Data Policy available at http://www.esrc.ac.uk/_images/DataPolicy_tcm8-4594.pdf accessed 24 March, 2011

Economic and Social Research Council (2010) Research Data Policy, available at http://www.esrc.ac.uk/_images/Research_Data_Policy_2010_tcm8-4595.pdf accessed 08 March, 2011

Economic and Social Research Council (2010) Research Ethics Framework, http://www.esrc.ac.uk/_images/Framework_for_Research_Ethics_tcm8-4586.pdf accessed 08 March

Environmental Information Regulations 2004, available at http://www.legislation.gov.uk/uksi/2004/3391/contents/made

JISC (2011) Virtual Research Environment Programme available at http://www.jisc.ac.uk/whatwedo/programmes/vre.aspx accessed 24 March, 2011

Rural Economy and Land Use Programme/UK Data Archive, (2010) Data Management Plan, available at http://relu.data-archive.ac.uk/plan.asp accessed 24 March, 2011

UK Data Archive (2011) Create and Manage Data, available at http://www.data-archive.ac.uk/create-manage accessed 24 March, 2011

UK Data Archive, Data Management Planning for ESRC Centres and Programmes, available at www.data-archive.ac.uk/create-manage/projects/JISC-DMP accessed 24 March, 2011

8 Appendices

8.1 List of dissemination activities

Date	Dissemination activity	Location	Description	Link
	UKDA_RDMSS Twitter	-	Twitter feed providing and linking to presentations of project progress and data management information	twitter.com/UKDA_RDMSS
	DMP-ESRC Facebook	-	Facebook page reporting on project progress and interaction with other data management actors	www.facebook.com/#!/pages/Colchester-United-Kingdom/Data-Management-Planning-for-ESRC-Research-Data-Rich-Investments/273698049591
	DMP-ESRC Blog	-	Blog updates on the progress of the project and other data management issues	ukdaresearchdatamanagement.blogspot.com/
	DMP-ESRC Webpage	-	Project web page containing information on the project aims and structure as well as outputs	www.data-archive.ac.uk/create-manage/projects/jisc-dmp
10-11/3/2010	JISCMRD Research Data Management Forum	Manchester	Presentation	
12/4/2010	New Dynamics of Ageing Programme Meeting on Maximising the Impact of Research	Birmingham	Presentation on DMP-ESRC project	www.data-archive.ac.uk/news-events/events.aspx?id=2757
17-18/5/2010	JISCMRD Programme Progress Workshop	Manchester	Presentation on DMP-ESRC project	www.esds.ac.uk/news/eventsdocs/DMPworkshop17may2010.pdf
1-4/6/2010	IASSIST 2010 Conference: Social Data and Social Networking: Connecting Social Science Communities Across the Globe.	Cornell University, USA	Presentation titled Data Management - Engaging Researchers and Crossing Disciplines - Experiences at the UK Data Archive	www.esds.ac.uk/news/eventsdocs/IASSIST2010vvde.pdf
June 2010	Evaluation panel JISC Grant Call 04/10 - Managing Research Data Training	-	Evaluating and marking proposals for the JISC Managing Research Data Training strand	http://www.jisc.ac.uk/whatwedo/programmes/mrd/rdmtrain.aspx
1/9/2010	UK Data Archive Newsletter and Website	-	Article: Archive Project Gets to the Heart of Good Data Management	www.data-archive.ac.uk/media/223131/ukdataarchivenews13.pdf
27/9/2010	New Dynamics of Ageing TACT3 project training event	Brunel University, London	Presentation titled: Data Archiving Explained	www.esds.ac.uk/news/eventsdocs/tact327oct2010.pdf
21/10/2010	PEKin Developing Data Management expertise in Research workshop	Kings College, London	Presentation titled Data Management - Support for Researchers	www.esds.ac.uk/news/eventsdocs/Developing_data_management_21oct2010.pdf
22/10/2010	JISCMRD Stage Two Launch Workshop	London	Participating in workshop, "Lift" pitch	
2/11/2010	JISCMRD Programme Costs, Benefits and Sustainability Workshop	Bristol	Presentation on costing activities for DMP-ESRC project	
19/1/2011	Incremental Project data management seminar	Cambridge University	Personal Data, Public Knowledge and Research Ethics	www.crassh.cam.ac.uk/events/1529/
19/1/2011	Review of DataTrain course outline	Cambridge University	Meeting with DataTrain project on Anthropology and Archaeology	www.lib.cam.ac.uk/preservation/datatrain/
15/2/2011	Data Management in Longitudinal Qualitative Research using Nvivo	Third Sector Research Centre, Univ. of Birmingham	Workshop on using Nvivo as a data management tool in a qualitative longitudinal research project	

18/3/2011	Managing research data: a growing priority, a fast-moving discipline	ESDS News Issue 5, March 2011 p.2	ESDS news article outlining projects run by the Archive's Research Data Management Support Service section	www.esds.ac.uk/news/newsdetail.asp?id=2824
28-29/3/2011	JISC Managing Research Data (International) programme workshop	Aston Business School Conference Centre, Birmingham	Attended and presented on DMP-ESRC project and findings	www.data-archive.ac.uk/news-events/events.aspx?id=2794
April 2011	IASSIST web resources on data management planning	-	Contributed referenced to data management planning resources, guides, examples and templates to the Data Management and Curation Resources collection on the IASSIST website.	http://www.iassistdata.org/resources/category/data-management-and-curation
4/5/2011	Data management planning and practices for Research Centres and Programmes	Royal Statistical Society, London	Seminar bringing together researchers, directors and support staff from ESRC-funded Research Centres and Programmes with funding councils and data services	http://www.data-archive.ac.uk/news-events/events.aspx?id=2807
5-6/5/2011	Research Data Management Forum VI: Planning for research data management	Leicester University	The new Managing and Sharing Data guide was presented and distributed at this event and we contributed to discussions on data management planning strategies for institutions and research funders	http://www.dcc.ac.uk/events/research-data-management-forum/rdmf6-planning-research-data-management
16/5/2011	Innogen data management planning recommendations and Q&A session for researchers.	York Marriott Hotel	Data management planning recommendations and Q&A session for researchers during Innogen's Research Programme Retreat.	
3/6/2011	IASSIST 2011 Conference: Data Science Professionals: a Global Community of Sharing	Simon Fraser University, Vancouver	Joint presentation with DCC, titled: Of Policy, Practice and Tools: Data Management Planning in the Social Sciences in the UK	http://www.rdl.sfu.ca/IASSIST/index.php/Program/f1

8.2 Questionnaire used during data management interviews

The purpose of this inventory is to understand research data in your project. We are not interested in the theoretical aspect of your research or your research question, just how you manage the collection and storage of data. Understanding this will help strategic planning and data management in this and other ESRC centres and programmes, and ultimately the wider research community. Please note this is not an audit or critical inspection of your operating procedures, so feel free to be frank about existing or potential problems or issues. Please indicate at any point where you feel you need further information and/or would like training on an aspect of data management.

Data management can be loosely defined as "Anything outside of actually analysing the data." So we are interested in the collection, organisation, protection, distribution and storage of data.

If you agree to take part in the project this will include being interviewed and audio recorded. The data we gather in this interview will be for internal information only, confidential, and not kept beyond the project.

Name of project:		
Principal Investigator(s)		
Your research role and responsibilities (please describe)		
Do you currently have a data management plan for your research?		
Yes		
No		
If yes, what was the motivator for devising/adopting one?	Research requirement to access/analyse. Annotate other's data	
	Requirement of the Centre	
	Requirement of the project funder	
	Size of the project team	
	Volume of data the project generates	
	Complexity of dealing with multiple formats	
	Other (Please specify)	
If no, what was the reason	Not required/appropriate to the field of research	

	Not a requirement of the Centre	
	Not a requirement of the project funder	
	Resources (time and effort) required	
	Lack of training/expertise	
	Lack of support/guidance	
	Don't know	
	Other (Please specify)	
Sources of data		
Created research data		
Secondary data		
Please state who owns copyright/Intellectual property rights for the data		
Principal Investigator(s)		
Centre or Programme		
Other (please specify). Also, please state under what conditions you accessed these data (licence, etc.)		
Who is funding your research (please select all that apply)		
ESRC		
Other (please specify). Please also state who asserts ownership of data.		
Please describe the formal process of ethical review for your research and any informal		
Please describe the process of obtaining informed consent from participants. If you have a consent form, please can you provide a copy.		
Does your data require anonymisation?		
Yes		
No		
If yes, please describe anonymisation procedures.		

What people are responsible for managing this data?	
You	
Research Project Manager	
Research Assistant	
Research Technician	
Graduate Student	
Other person in Centre	
Departmental IT Officer	
University IT department	
External Data Centre	
Data archive	
Don't know	
No one	
Other (please specify)	

Do you plan to make your data available for reuse/sharing?		
Yes		
No		
Partly		
If yes or partly, through what means? Please describe.		
If no or partly, what reasons?	Confidentiality/Data protection issues	
	Licence agreements	
	The data is not fully documented for reuse/sharing	
	Lack of resources to prepare data and documentation	
	The data is in a format that is not widely machine readable	
	Other (please specify)	

Changes to and/or updates of data and documentation	Documentation	Data
Daily		
Weekly		
Monthly		
Yearly		
Reference only		
Other (please specify)		

Time associated with creating data	

	Fieldwork	Data entry/transcription
Hours		
Days		
Weeks		
Months		
Years		
Other (please specify)		

Please can you estimate the cost of data creation?

Methods of primary data collection	
N/A	
Survey	
Interview	
Focus group	
Observation	
Database compiled from documentary sources	
Other (please specify)	

What quality assurance procedures are in place for data and documentation?		
	Documentation	Data
Securing master copies to avoid tampering		
Version control of amended data/documentation		
Transcribing procedures and transcription template		
Labelling of primary data and metadata		
Explanation of non-standard acronyms		
Data validation and error checking procedures (i.e. two pass verification, range checking). Please describe.		

What type of documentation is produced?	
User manual	
Codebook	
Interview schedule	
Questionnaire	
Data list (i.e. Interview information)	

Please estimate the size of your research data

<1 GB	
1-50 GB	
50-100 GB	
100-500 GB	
500 GB - 1 TB	
1 - 50 TB	
50 - 100 TB	
> 100 TB	
Don't know	

Do you currently allow others outside the research team to access your data?		
If yes, who?	Students/colleagues in department	
	Students/colleagues in Centre	
	Colleagues at other institutions	
	As supporting evidence	
	General public	
	Other (Please specify)	
If no, what access issues are of concern to you?	Confidentiality of data	
	Intellectual property rights	
	Commercial value of data	
	Possible misinterpretation of data	
	Time/effort required	
	Other (Please specify)	

Are any critical research data assets stored on proprietary formats? Please describe

Please indicate how your data is collected, stored and if applicable, transmitted			
	Collected	Stored	Transmitted
Hard disk of data recording instrument (i.e. on voice recorder or camera)			
Computer hard disk connected to a network/internet			
Computer hard disk not connected to a network/internet			
External hard drive			
Local server			
University server			
Third party/commercial server			

CD/DVD			
USB/Flash device			
Email			
Hard copy (paper)			
Other (please specify)			
How is your data secured?			
No security measures			
Password protected			
Encrypted			
Other (please describe), this may include physical security measures (i.e. locked filing cabinets or safes)			
Is your data backed up?			
Yes			
No			
If yes, how frequently?	Daily		
	Weekly		
	Monthly		
	Ad hoc		
	Don't know		
	Other (please specify)		
What is backed up?	Everything		
	Data critical to the project		
	Data required for		
	Don't know		
	Other (please specify)		
Where is it backed up?	Hard disk of data recording instrument (i.e. on voice recorder or		
	Computer hard disk connected to a network/internet		
	Computer hard disk not connected to a network/internet		
	External hard drive		
	Local server		
	University server		
	Third party/commercial server		
	CD/DVD		
	USB/Flash device		
	Paper (hard copy)		

	Other (please specify)		
Have you encountered storage problems?			
Yes			
No			
Not applicable			
If yes, how?	Insufficient storage space		
	Files corrupted		
	Lost/missing storage medium		
	Other (please specify)		
When creating or using your data, did you collaborate with anyone else?			
Yes			
No			
Did you encounter practical problems with data sharing when collaborating?			
Yes			
No			
Not applicable			
If yes, what?	Finding suitable shared storage space		
	Lack of file naming conventions made it difficult to identify files		
	Lack of version control made it difficult to identify most recent version		
	Other (please specify)		
Do you have a planned retention period for data and documentation you have created?			
Yes			
No			
Not applicable			
If yes, how long?		Documentation	Data
	< 1 year		
	1-2 years		
	2-5 years		
	5-10 years		
	11 years +		
	Indefinite		
	Don't know		
	Other (Please specify)		

If no, why not?	Not a formal requirement		
	Don't know		
At either the end of the research or retention period, have you been asked to destroy data and/or documentation?			
		Documentation	Data
Yes			
No			
Partly			
If yes or partly, please indicate how	Deleting data		
	Overwriting software		
	Physical destruction		
Thinking back over the areas covered in this interview, are there any areas where you feel you need further information and/or would like training on an aspect of data management?			

Thank you for completing this survey, if you have any questions or comments please contact the project Research Data Management Senior Officer, Laurence Horton: lhorto@essex.ac.uk