Survey on Research Data Management and Sharing: Working Paper

Title: Research Data Management Web Survey – working paper

Project Name: Research Data Management Support Service (RDMSS)

Work Package: 4.2

Draft/Version: Version 1

Catalogue Information

Title:	Research Data Management Web Survey: Working Paper
Creator:	Gareth Knight
Keywords:	data management, web survey, research data
Description:	This working paper describes the web survey on research data management performed at the London School of Hygiene & Tropical Medicine
Publisher:	London School of Hygiene and Tropical Medicine
Date:	2012.10.01
Language:	English
Rights:	Creative Commons Attribution 3.0 Unported License

Report Version Control

Version	Date	Author	Change Description
Draft 1	18 Sep 2012	Gareth Knight	First draft for comment
Draft 2	28 Sep 2012	Gareth Knight	Ongoing revision based upon comment provided during the trialing stage
Version 1.0	03 Oct 2012	Gareth Knight	Ongoing revision based upon comment from various stakeholders. Removal of five questions to reduce the amount of time that people will spend completing the survey.

1. Introduction

1.1 Overview

This document sets out a question list to be applied in a web survey performed on the topic of research data management activities within the London School of Hygiene and Tropical Medicine.

1.2 Structure of the Survey

The survey is comprised of four pages in total:

1. **Cover Page:** The cover page provides basic information on the survey, indicating its purpose and the data controller responsible for its construction. Sample text for the cover page is provided on p3

2. **Survey Page:** The second page contains questions 1-9. These questions are intended to gather information about the researcher, the research project, the information contained within the research data, and the approach that is taken to its short-term storage. A summary of the question list is provided on p3-4

3. **Survey Page:** The third page contains questions 10-18. These questions are intended to gather information about the IPR status of the research data, sharing long-term management requirements, sharing requirements, and training needs. A summary of the question list is provided on p3-4.

4. **Acknowledgement Page:** The final page provides an acknowledgement and contact details. The survey is comprised of twenty four questions in total, of which twelve are mandatory. Sample text for the acknowledgement page is provided on p6.

The current version of the survey is available at the following URL: https://www.survey.lshtm.ac.uk/lshtm-rdm/

1.2.1   Cover Page

Research Data Management Survey
Welcome to the Research Data Management Survey 2012

We would like to gain a better understanding of the type of digital research data that you produce and use, and the areas of support required to help you to manage your research data in compliance with funding obligations and in-line with best practice.

Responding to the survey will help us to us to plan and prioritise the support that should be provided to researchers within the School, including training, guidance, and technical systems. The survey consists of 15 questions and should take you approximately 10-15 minutes to complete.

What is Research Data?
For the purpose of this survey, Research Data refers to any type of data created, collected, or generated in a digital form that is analysed to produce original research results.

Who should complete this survey?
The survey is open to anyone who is actively involved in funded or unfunded research. This may include Principal Investigator of a research project, Data Managers, and research staff involved in the process.

Confidentiality
All data collected in this survey will be held anonymously, securely and in strict confidence. Responses are not linked back to individuals, and results will be presented in an anonymised form. If, however, you would like to discuss your data storage, management, and sharing requirements with a member of the RDM Support Service, please enter your email address at the end of the survey.

The data controller for all information collected in this survey is Gareth Knight, RDM Support Service Manager - telephone 020 7927 2564, or email gareth.knight@lshtm.ac.uk.

Note that once you have clicked on the CONTINUE button at the bottom of each page you can not return to review or amend that page.

1.2.2   Survey Pages

No.	Question	Obligation	Response Type
1	Which faculty are you located in?	mandatory	Single response
2	Who, if anyone, funds the research project (or projects) that you are currently undertaking?	Optional	Multiple response
3	Please describe your research data	Mandatory	Multiple response
4	What are the three most common software products that you use to create and manipulate your research data? (e.g. STATA, MySQL)	Optional	Unstructured text
5	Where do you store your data during the life of the project/ research activity?	Mandatory	Multiple response
6	How many years are you required to retain the Research Data for?	Mandatory	Single response
7a	Are there legal, regulatory, or confidentiality issues that influence how your Research Data is stored, managed	Mandatory	Single response

	and/or shared?		
7b	If yes, please indicate the legislation, policies, or other rules that applies to your data.	Optional	Multiple response
8a	Do you currently apply any type of security to all or some of the Research Data produced by your project?	Mandatory	Single response
8b	If yes, please indicate the security measure(s) that you currently apply.	Optional	Multiple response
9a	Do you share the Research Data that you create/manage beyond the project team during the life of a project / research activity?	Mandatory	Single response
9b	Who do you typically share some or all of the Research Data with during the life of a project / research activity?	Mandatory	Multiple response
10	What arrangements will be made to provide access to research data following completion of the funded project/unfunded research activity? E.g. in an anonymised form.	Mandatory	Single response
11a	What issues or challenges have you encountered when creating, managing, and/or sharing your research data?	Optional	Multiple response
11b	Please provide further details on the issue/challenge	Optional	Unstructured text
12	Please rate your interest in training on the following topics related to data management?	Mandatory	Likert Scale
13	What other training related to data management do you require?	Optional	Unstructured text
14	What resources should the School provide to help you to manage your Research Data? (e.g. increased capacity on network storage, provision of a data archive, etc.)	Optional	Unstructured text
15	Please enter your email address if you would like to discuss data management and sharing needs within your own project	Optional	Email address

1.2.3 Acknowledgement Page

Research Data Management Survey
Thank you for participating.

Findings from the survey will be made available in due course.

If you would like to contact the RDM Support Service to discuss your data management and sharing needs, please use the following address:

Telephone: 020 7636 2564
Email: researchdatamanagement@lshtm.ac.uk

Gareth Knight
RDM Support Service Manager

2. Survey Questions

2.1 About You

The purpose of the first section is to gather information on the respondent themselves.

2.1.2. Question 1

	Description
No.	1
Question Text	Which faculty are you located in?
Description	Establishes the Faculty in which the respondent is located within the School
Purpose	• The respondent's faculty may provide context in understanding their subsequent responses • In combination with Q9, it may provide insight into the technical infrastructure & services available to researchers working in each Faculty • In combination with Q12 and 16, may be used to identify specific challenges that are encountered by researchers in a specific faculty • In combination with Q17, may be used to identify trends in training needs across each faculty • In combination with Q18 and Q19, may be used to identify future needs of each faculty
Obligation	Mandatory
Response Type	Single response
Representation	Option box + free text field for other
Responses	• Faculty of Epidemiology and Population Health • Faculty of Infectious and Tropical Diseases • Faculty of Public Health and Policy • Central services (IT, library and archives service, research office, etc.) • Division of Education • Prefer not to say • Other (please state)

2.2 About Your Research

The second section contains questions related to the research activities performed by the respondent. The respondent is asked to select one or more funded or unfunded research project that they are currently working upon and complete questions 3 - 11 with reference to the work.

2.2.1. Question 2

	Description
No.	2
Question Text	Who, if anyone, funds the research project (or projects) that you are currently undertaking?
Description	Identifies one or more agencies that fund the research project
Purpose	• Establishes requirements and constraints for data management and sharing that are imposed upon the researcher. • May be used to determine if respondent is aware of funder data management and sharing requirements
Obligation	Mandatory
Response	Multiple response

Type	
Representation	Checkbox + free text field for other
Responses	• Biotechnology & Biosciences Research Council (BBSRC), • Bill & Melinda Gates Foundation, • Cancer Research UK (CRUK) • Department of Health UK (DoH) • Dept for International Development (DID) • Drugs for Neglected Diseases Initiative (DNDI) • Economic and Social Research Council (ESRC) • GlaxoSmith Kline (GSK) • Medical Research Council (MRC) • Wellcome Trust • World Health Organisation (WHO) • World Cancer Research Fund (WCRF) • Other (please specify):
Notes	

2.3   Your Research Data

The third section contains questions related to the research data produced and/or used by the project.

2.3.1.   Question 3

	Description
No	3
Question Text	Please describe your research data
Description	A description of the type of research data created and/or used within the project and the form in which it is stored.
Purpose	• May be used to determine if the research data requires curation and preservation and the appropriate data archive in which it should be held • May be used to evaluate the suitability of security arrangements
Obligation	Mandatory
Response Type	3 x single response

Representation	Three sub-questions are asked:
	   a. Does it contain primary data created during your research?
	   b. Does it contain quantitative data?
	   c. Does it contain personally identifiable information at any stage of the lifecycle? (e.g. prior to anonymisation)

Responses	Each sub-question may be addressed using the following: Yes No To some extent I don't know Prefer not to say Not applicable
Notes	The question has been designed around the limitations of the grid question type in BOS tool, hence the use of 'does X contain Y' style questions that have a Y/N/other response. Guidance is provided indicating that 'To some extent' indicates that data contains a combination of different data (e.g. quantitative and qualitative, primary and secondary)

2.3.2. Question 4

	Description
No	4
Question Text	What are the three most common software products that you use to create and manipulate your research data? (e.g. STATA, MySQL, MS Excel)
Description	A description of the software packages that the respondent most commonly uses to collect and/or analyse research data.
Purpose	• Identifying the software tools that researchers use to create and/or analyse research data will provide insight into their working environment, the file formats that they are likely to use, the potential problems that may be encountered, and the type of guidance that needs to be provided (e.g. if they use a SQL database, ensure that guidance is provided on DB lock and preservation formats) • Responses may contribute to ongoing discussion on technical infrastructure that should be provided by the School to support researchers
Obligation	Mandatory
Response Type	Unstructured text
Representation	3 x free text fields (Tool A, Tool B, Tool C)
Responses	-
Notes	Respondents are advised to enter 'None' if they have 1 or 2 preferred tools.

2.3.3. Question 5

	Description
No	5
Question Text	Where do you store your data during the life of your project/ research activity?
Description	Identifies the location that is commonly used to store research data
Purpose	• Provides information on the facilities that the respondent uses to store their Research Data • May be used to infer the number of different locations used to store research data (use of 1 storage system is a potential risk)
Obligation	Mandatory
Response Type	Multiple response
Representation	Checkbox + free text
Responses	• Local drive on my School computer, • Local disk drive on a laptop/netbook, • Portable storage device (e.g. external drive, USB disk), • Network servers dedicated to the project at School (NAS Box/ PC shared drive) • Network storage system/servers maintained by collaborating institution (NAS Box/ PC shared drive) • Home/H: drive of the School servers • Shared storage area of the School servers • CD/DVD • Email system • Content/data management system operated by the project or School • Content/data management system operated by a project partner/collaborator • School iFolder service • Web based service, (e.g. Dropbox, Flickr, Google Docs)

	• Other (please specify)
Notes	

2.3.4. Question 6

	Description
No	6
Question Text	How many years are you required to retain the Research Data for after the end of the project?
Description	The number of years that the research data must be retained to meet funder, legislative, and/or other requirements.
Purpose	• Establishes respondent's awareness and understanding of funder expectations for data management. A response that differs from that provided by the funder or is not completed would suggest that they have a poor understanding of retention requirements. • May be used to measure compliance with funder, legislative and/or requirements, in combination with Q3
Obligation	Mandatory
Response Type	Single response + free text (for other)
Representation	Drop-down list + free text
Responses	• 1 – 5 years • 6 – 10 years • 11 – 15 years • 16 – 20 years • 20+ years • Varies – influenced by date of last access • Indefinitely • I don't know • Other
Notes	

2.4 Data Security

The fourth section contains questions related to the management of Research Data with relations to legal compliance.

2.4.1. Question 7

	Description
No	7a
Question Text	Are there legal, regulatory, or confidentiality issues that influence how your Research Data is stored, managed and/or shared?
Description	An indicator of the requirements and regulations that must be met when managing research data
Purpose	• Establishes respondent's awareness and understanding of the requirements and regulations that must be met. Used in combination with Q3
Obligation	Mandatory
Response Type	Single response
Representation	Option box
Responses	• Yes • No • I don't know • I prefer not to say

Notes	

	Description
No	7b
Question Text	If yes, please indicate the legislation, policies, or other rules that applies to your data.
Description	An indicator of the requirements and regulations that must be met when managing research data
Purpose	• Establish the legal, regulatory, or confidentiality issues that must be addressed by a data archive/data enclave.
Obligation	Mandatory
Response Type	Multiple response
Representation	Check box + free text
Responses	• Data Protection Act • Freedom of Information • Funder requirements • NHS ethics/NRES • Environmental Information Regulations (EIR) • School requirements (e.g. research governance) • Institutional policy of a third party (e.g. one or more collaborating institutions) • Other (please specify)
Notes	Although it is possible that some projects will produce data that must comply with EIR, its primary purpose is to establish whether the respondent is thinking about their choice or simply selecting every single bot.

2.4.2. Question 8

	Description
No	8a
Question Text	Do you currently apply any type of security to all or some of the Research Data that you create and/or use within your research?
Description	An indicator of the security measures that are applied to the research data, if any.
Purpose	• Establishes how the researcher addresses the obligations set out in their Q12b response.
Obligation	Mandatory
Response Type	Single response
Representation	Option box
Responses	• Yes • No • I don't know • I prefer not to say
Notes	

	Description
No	8b
Question Text	If yes, please indicate the security measure(s) that you currently apply.
Description	An indicator of one or more measures that are applied to the research data to address security obligation

Purpose	• Establish the legal, regulatory, or confidentiality issues that must be addressed by a data archive/data enclave.
Obligation	Optional
Response Type	Multiple response
Representation	Check box + free text
Responses	• Encryption • Anonymisation • Password protection of files • Access logging • Physical security (e.g. locked room) • Other (please specify)
Notes	

2.5 Data Sharing

The fifth section addresses activities performed by the respondent to share research data with other stakeholders during the life of the project and following its completion.

2.5.1. Question 9

	Description
No	9a
Question Text	Do you share the Research Data that you create/manage beyond the project team during the life of a project / research activity?
Description	One or more stakeholders with whom the research data will be shared during the lifetime of the funded project / unfunded research activity
Purpose	• Establishes the short-term requirements for data sharing – information that may prove useful if the RDMSS project adopts a data management system that manages data during the project lifetime.
Obligation	Mandatory
Response Type	Single response
Representation	Option box
Responses	• Yes • No • Varies between projects • I don't know • I prefer not to say
Notes	

	Description
No	9b
Question Text	Who do you typically share some or all of the Research Data with during the life of the project / research activity?
Description	One or more stakeholders with whom the research data will be shared during the lifetime of the funded project / unfunded research activity
Purpose	• Establishes the short-term requirements for data sharing – information that may prove useful if the RDMSS project adopts a data management system that manages data during the project lifetime.
Obligation	Mandatory
Response Type	Multiple response

Representation	Checkbox + free text
Responses	• Third party Data Provider / Data Creator • One or more members of the research group • One or more members of the department • Collaborating partners at other institutions • Funder(s) • Publisher(s) • Anyone who expresses an interest • Anyone within the School • Wider public • Other (please specify)
Notes	

2.5.2. Question 10

	Description
No	10
Question Text	What arrangements will be made to provide access to research data following completion of the funded project/unfunded research activity? E.g. in an anonymised form.
Description	Actions that will be performed to manage research data and make it available following the project's completion
Purpose	• May be used to gauge the respondent's understanding of obligations and commitment to maintaining research data over time, when used in combination with Q3. • Indicate the approach that will be adopted to manage research data during the post-funding phase, which may inform the scoping activities for the data management system (e.g. data that should be managed, systems with which it must interact).
Obligation	Mandatory
Response Type	Single response
Representation	Option box + free text
Responses	• Data will be deposited with a third party data service or archive (e.g. the Economic and Social Data Service) and made available for access • Data will be made available through the project or other website (e.g. subject to user registration) • Data will be held within the department or project team. A request may be submitted to a Research committee or project team to gain access to research data • Data cannot be made publicly available. Access will be limited to the PI, project team, or other designated individuals • I do not know • Other (please specify)
Notes	

2.5.3. Question 11

	Description
No	11a
Question Text	What issues or challenges have you encountered when creating, managing, and/or sharing your research data?
Description	Issues, challenges, problems encountered when creating, managing, and/or sharing research data
Purpose	• Identify problems that researchers encounter when handling

	research data, which may be used to inform development work in the RDMSS project and/or other departments/units within the School.
Obligation	Mandatory
Response Type	Multiple response
Representation	Checkbox + free text
Responses	• Lack of storage space on School network • Speed of access to data • Uncertainty on file formats to use • Uncertainty on documentation standards to apply • Issues when preparing a data sharing agreement • Issues when producing a data management plan • Security issues • Uncertainty on practices related to data archiving • Interoperability issues (e.g. between different software applications, operating systems) • Other (please specify)
Notes	

	Description
No	11b
Question Text	Please provide further details on the issue/challenge
Description	Issues, challenges, problems encountered when creating, managing, and/or sharing research data
Purpose	• Identify problems that researchers encounter when handling research data, which may be used to inform development work in the RDMSS project and/or other departments/units within the School.
Obligation	Optional
Response Type	Unstructured text
Representation	Free text
Responses	-
Notes	

2.6 Future Data Management Support

The sixth and final section is intended to gauge the respondents' thoughts on the support that should be provided within the School to help them to manage and share research data.

2.6.1. Question 12

	Description
No	12
Question Text	Please rate your interest in training on the following topics related to data management
Description	An indication of the training needs that should be provided by the School
Purpose	• An indication of the data management training that should be provided by the School. Will inform development of the education programme during Stage 2 (January 2013 onwards)
Obligation	Mandatory
Response Type	Likert scale
Representation	9 option boxes
Responses	The respondent will be asked to rate the following training course using one of the provided ratings

1 = Not required
2 = May be interested
3 = Definitely interested
4 = No opinion

Training course:
- Developing a research data management plan
- Documenting your data
- Storing your data
- Sharing your data
- Creating metadata for data
- Ethics and consent
- Funders requirements and research data management
- Copyright and Intellectual Property Right (IPR)
- Citing your Research Data

Notes	

2.6.2. Question 13

	Description
No	13
Question Text	What other training related to data management do you require?
Description	An indication of the training needs that should be provided by the School
Purpose	• An indication of the data management training that should be provided by the School. Will inform development of the education programme during Stage 2 (January 2013 onwards)
Obligation	Optional
Response Type	Unstructured text
Representation	Free text
Responses	-
Notes	

2.6.3. Question 14

	Description
No	14
Question Text	What resources should the School provide to help you to manage your Research Data? (e.g. increased capacity on network storage, provision of a data archive, etc.)
Description	Resources that the School should provide to enable the researcher to perform their work more effectively and easier
Purpose	• Establish future infrastructure & development needs
Obligation	Optional
Response Type	Unstructured text
Representation	Free text
Responses	-
Notes	

2.6.4. Question 15

	Description
No	15
Question Text	Please enter your email address if you would like to discuss data management and sharing needs within your own project
Description	Please provide your contact details
Purpose	• Identifies respondents who are willing to be interviewed.
Obligation	Optional
Response Type	Unstructured text
Representation	Free text
Responses	-
Notes	