



OSLO AND AKERSHUS  
UNIVERSITY COLLEGE  
OF APPLIED SCIENCES



Ewelina Melnarowicz

---

**Because research data matters –  
a case at the Loughborough University Library**

Supervisor: Pat Dixon

## Abstract

Research data management is increasingly gaining importance in academic libraries, and there is a growing body of literature dedicated to this area of interest. However, only a few studies explored opinions and requirements of stakeholders involved in the provision of research data management.

The following research employed a qualitative case study method and the evidence was gained through semi-structured interviews and document analysis. It aimed to investigate research data practices at Loughborough University, explore how research data management will be implemented investigating roles, responsibilities and work practices, evaluate the outcome of activities taken in the implementation of research data management and finally analyse whether/how users are going to be involved in this implementation.

The study determined opinions and requirements of stakeholders involved in the provision of research data management and, based on the evidence, identified and discussed five main categories responsive to the objectives of the study: current services, stakeholders and roles, current issues, future developments and solutions. It also outlined recommendations for Loughborough University.

The study confirmed that there are no proven solutions for the management of research data, therefore, further research is suggested to respond to the resulting questions that emerged from this study.

## Table of contents

Acknowledgments .....	VI
Declaration and Plagiarism Disclaimer.....	VII
List of abbreviations and acronyms.....	VIII
CHAPTER 1: Introduction .....	1
1.1 Background information .....	1
1.2 Statement of the problem.....	3
1.3 Research question.....	3
1.4 Aim and objectives .....	3
1.5 Methodology .....	4
1.6 Scope .....	4
1.7 Limitations .....	4
1.8 Significance of the research .....	5
1.9 Structure of the thesis.....	5
CHAPTER 2: Loughborough University .....	7
2.1 The Loughborough University Library .....	8
CHAPTER 3: Literature review.....	10
3.1 Introduction .....	10
3.2 Definition of research data management.....	10
3.3 Drivers of research data management.....	13
3.3.1 Data deluge.....	13
3.3.2 Research practice .....	15
3.3.3 Research councils.....	16
3.3.4 Academic libraries .....	17
3.4 Implementation of research data management.....	22
CHAPTER 4: Methodology .....	25
4.1 Research paradigm and approach .....	25
4.2 Research method.....	26
4.3 Sample and sampling methods.....	26
4.4 Data collection techniques.....	28
4.4.1 Interviews.....	28
4.4.2 Document analysis .....	31

4.5	Data analysis .....	31
4.6	Constrains .....	32
4.6.1	Limitations in methods, data collection and analysis .....	32
4.7	Ethical considerations .....	34
4.8	Trustworthiness of research – Value of research .....	34
4.8.1	Credibility .....	35
4.8.2	Transferability .....	35
4.8.3	Dependability .....	35
4.8.4	Confirmability .....	35
CHAPTER 5: Data analysis and discussion .....		37
5.1	Data Analysis .....	37
5.1.1	Current activities .....	37
5.1.2	Stakeholders and roles.....	41
5.1.3	Current issues .....	47
5.1.4	Future developments .....	50
5.1.5	Solutions .....	54
5.1.6	Value of managing research data .....	56
5.2	Discussion.....	57
5.2.1	Current activities .....	57
5.2.2	Stakeholders and roles.....	59
5.2.3	Current issues .....	61
5.2.4	Future developments .....	63
5.2.5	Solutions .....	64
CHAPTER 6: Conclusions and implications.....		66
6.1	Research Objectives .....	66
6.1.1	Examine how research data practices are currently supported .....	66
6.1.2	Explore how RDM support and services will be implemented, including roles and responsibilities, work practices .....	69
6.1.3	Evaluate the outcome of activities taken in the implementation of RDM .....	71
6.1.4	Analyse whether/how users are going to be involved in the implementation of RDM.....	74
6.2	Recommendations for Loughborough University and the University Library .....	76
6.3	Implications for further research .....	76

References ..... 78

Appendices ..... 89

    Appendix 1: Introductory email ..... 89

    Appendix 2: Participant Information Sheet ..... 90

    Appendix 3: Informed consent form ..... 91

    Appendix 4: Interview protocol ..... 92

    Appendix 5: Email interview ..... 93

    Appendix 6: Interview guide ..... 94

## **Acknowledgments**

This thesis would have not been concluded without the support of the following people:

My most sincere gratitude goes to the interviewees who shared their precious knowledge and time. All my professors that contributed to my education and who made this course invaluable, especially Pat for her support and advice. All my DILL friends for the fabulous moments we spent together and their enthusiasm. My heartiest thanks go to my family, for their encouragement, and to my dearest Aldo who became a full time house manager for the last two years.

**Declaration and Plagiarism Disclaimer**

“The opinions expressed in this dissertation are solely those of the author and acceptance of the dissertation as a contribution to the award of a degree cannot be regarded as constituting approval of all of its contents by the Division of Information & Communication Studies”.

I certify that all material in this dissertation which is not my own work has been identified and properly attributed.

Signed:.....*Qmelnowicz*.....

Date:.....*27/06/2012*.....

## List of abbreviations and acronyms

CMS	Content Management System
DCC	Digital Curation Centre
EPSRC	Engineering and Physical Science Research Council
FTE	Full Time Equivalent
HPC	High Performance Computing
LU	Loughborough University
RDM	Research Data Management
UK	United Kingdom
USA	United States of America



# CHAPTER 1: Introduction

## 1.1 Background information

The importance of managing digital materials, including research data, has been broadly recognised, and there is a variety of institutions and activities dedicated to this topic including the Digital Curation Centre in the UK, the International Journal of Digital Curation and the Digital Preservation Coalition. In particular the management of research data has gained interest in the last decades (ARL/NSF Workshop, 2006, p. 16). According to a succinct definition provided by Yakei “digital curation is the active involvement of information professionals in the management, including the preservation, of digital data for future use.” (Yakei, 2007). Furthermore, Beagrie concluded that digital curation is central not only to making use and accessing digital assets but also to our digital knowledge: “For society and individuals, it can be argued that digital knowledge if it is to be useful and useable must be continuously updated, maintained, and accessed. The emerging field of digital curation is central to this process.” (2006, pp. 12–13)

Prior studies have defined data as the foundation of research (Hey & Trefethen, 2003; Lord & Macdonald, 2003) or a critical part of it (Lynch & Carleton, 2009) and highlighted the need to manage research data along with other research outputs. “Increasingly, data are being recognized as first-class intellectual objects that can undergo quality checks, peer review, distribution, and reuse. The reuse of data contributes as much to society as the reuse of a concept in a journal article.” (Heidorn, 2011). In fact, it has been acknowledged that the research process is enhanced by managing and sharing of data, as the results gain transparency and because it allows to realise the full value of data (2011, p. iii).

Research funders also acknowledged this care for scholarly outputs, in particular in the USA and UK, and have been increasingly placing expectations on researchers to include a data management plan and share research results. According to Van den Eyden et al. (2011, p. 1) this requirement is now among the funders’ priorities, and in the UK all the Research Councils and other research bodies now mandate the deposit and sharing of data, which is also reflected in their policies.

Traditionally, academic libraries have provided research support and have been involved in managing, preserving and providing access to information. The focus in this study is on academic libraries and especially on research libraries, which are research led and are affected by some of the major transformations in the research practice (Lynch & Carleton, 2009). Nowadays their wide remit includes delivering a variety of print and digital collections and teaching through courses, workshops or face to face meetings.

Whilst academic libraries have primarily focused their activities on traditional research outcomes, such as journal articles, books or various artefacts of research, increasingly research led libraries support data intensive research.

Some authors claimed a certain enthusiasm of librarians, who engaged in data curation. “While libraries identified a significant list of pressure points, an overall enthusiasm for new roles in the academic research process was evident throughout the survey responses and in the case studies.” (Soehner, Steeves, & Ward, 2010). However, this issue was not only described as unfamiliar or daunting (Haas & Murphy, 2009; Peters & Dryden, 2011), but, according to Gold, it is a cultural shift (Gold, 2010) involving values of professionals that support research. “At the most fundamental level, engaging the library profession in the problem of data management may lead to reframing the values and practices of the library profession” (Gold, 2007a). Similarly, Heidorn observed:

When academic library administrators first hear that scholarly data now fall within the purview of the library, they may lose many nights’ sleep wondering who has cast this curse upon them and what resources will be needed to actually manage the data responsibly. (Heidorn, 2011)

Although libraries are challenged by this changing scholarly practice, they see the opportunity to reinvent themselves and redefine their role to become a key factor in the research (Thomas, 2011). This has also been described with certain emphasis as “once-in-a-lifetime opportunity” for libraries and “librarians will have to embrace the role of data curator to remain relevant and vital to our scholars” (Ogburn, 2010, p. 243).

In order to tackle occurring changes in the research arena, various academic libraries have embarked on research data projects or set up new research support services. According to Gold, there are many viable ways for libraries to support scholars and contribute to the data universe (Gold, 2007a). Some libraries are delivering services related to various stages of the research process such as conceptualisation, discovery, funding or experimentation, ranging from advice on research data to taking part in the research team (Auckland, 2012, pp. 20–24). Other institutions gathered requirements through pilot studies and reported in details on the researchers’ needs (Jake Carlson, 2012; HATII & University of Glasgow, 2009; Wilson et al., 2010).

This study is related to a project that the researcher carried out while working as an intern at the Loughborough University Library. She was involved in a pilot study on research data management which aimed to ascertain some of the key requirements of researchers at Loughborough University. As part of the project, a range of researchers coming from different disciplines was interviewed to collect evidence on the key issues facing researchers when dealing with data. This project offered one perspective, those of researchers, on the research data and inspired this enquiry, which sought the perspective of other

stakeholders.

## **1.2 Statement of the problem**

In the context of digital curation, the debate around research data management is ongoing and, to some degree, it is still a new concept represented by a relatively few elite initiatives (Gold, 2010).

As stated above, there are several studies that explored needs and requirements of researchers and only few studies explored thoroughly research data management (Denison, Kethers, & McPhee, 2007; Martinez-Uribe, 2009) from the perspective of researchers and service providers. A study at Oxford University looked at requirements of researchers from the perspective of service providers by analysing services on offer and identifying available expertise. The findings revealed that representatives of service units needed guidance on how to provide these new services (Martinez-Uribe, 2009). In the study of Denison et al. (2007), the authors discussed researchers' requirements and, more specifically, some of the issues regarding the adoption and provision of data repositories. In this study individuals working in various institutional research support entities, information professionals and archivists were considered as important "systemic players" to get an insight into barriers and to understand local and external collaboration. However, the authors did not provide further discussion regarding these people.

It is apparent that evidence on opinions and requirements of some of the stakeholders involved in the provision of research data management is still scarce. Therefore, this study attempted to explore and analyse opinions of individuals involved in the management of research data, in particular librarians, research support and IT staff aiming to understand their opinions and requirements.

## **1.3 Research question**

This study will attempt to answer the following research question:

What are the requirements and opinions of some of the stakeholders involved in the provision of research data management?

Understanding of requirements and opinions of various stakeholders is relevant to enhance the provision of academic library services.

## **1.4 Aim and objectives**

The aim of this research was to gain an understanding of requirements and opinions of stakeholders involved in the provision of research data management at Loughborough University.

Accordingly, this dissertation sought to achieve the following objectives:

- Investigate how research data practices are currently supported
- Explore how research data management will be implemented, including roles, responsibilities and

work practices

- Evaluate the outcome of activities taken in the implementation of research data management
- Analyse how/whether users are going to be involved in the implementation of research data management

## **1.5 Methodology**

The study followed the line of research from a constructionist perspective and employed a qualitative case study method, in which the evidence was gathered from interviews and documents.

The sample approach occurred in two stages: in the initial plan, the choice of participants was made according to snowball sampling to identify the profiles of potentially information-rich individuals. In the second stage, according to purposive sampling and seeking to gather multiple perspectives, other individuals were identified and invited to participate in the study. The population of the study included all Loughborough University staff involved in the implementation of research data management, namely the Library, and especially the Library Research Support Group, the Research Office, High Performance Computing, Intellectual Property Office and IT Services. In total, eight members of Loughborough University participated in the interviews. They were on different positions and from different departments; the majority of them was affiliated with the library (4), and the other four participants were from the Research Office (2), from the Research Computing (1) and IT Services (1).

## **1.6 Scope**

This study focused on research data management initiated at the Loughborough University Library. More specifically, it is limited to some of the stakeholders involved in this process, namely the Library, the Research Office, IT Services and a selection of institutional documents, ranging from policies to minutes meetings.

## **1.7 Limitations**

Resources and time were the main constrains of this research. Other limitations are related to the choice of research methods and data techniques employed. Certainly, the case under study would have gained the fullest understanding if both quantitative and qualitative methods had been employed. Furthermore, additional data sources covering documents from a range of support services on campus not limited to the library would have been helpful to achieve a rich and detailed evidence of enquiry. These and other potential limitations are further discussed in Chapter Four.

## **1.8 Significance of the research**

The results of this research shall inform relevant stakeholders of research support within Loughborough University on current services and present the point of view and main requirements of librarians and other stakeholders, therefore, giving them an opportunity to enhance the service.

This study shall contribute to the existing research on the research data and academic libraries and will act as a source of reference for those aiming to do further research on this topic.

## **1.9 Structure of the thesis**

This thesis consists of six chapters. Beginning with an introduction to the topic of the study, Chapter One presents the research question, aim and specific objectives that the study aimed to address. It also defines the scope of research and outlines its main limitations that were faced during the research. Finally, it presents the significance of this thesis.

Chapter Two introduces the reader to Loughborough University and the University Library that provide the context of the case study.

Chapter Three is a review of the literature that discusses the key publications on the topic of research data management and, where necessary, brings in other related information. The review begins with a definition of research data management. Then, it examines the key drivers of research data management, covering data deluge, research practice, research funders and academic libraries. Finally, it illustrates the implementation of RDM.

Initially, it examines definitions of research data management and other related concepts, covering data curation, digital preservation, stewardship and archiving. The next section outlines the key drivers to research data management, including data deluge, research practice, research funders and academic libraries. It concludes by considering implementation of research data management in academic libraries.

Chapter Four presents the methodology – a qualitative case study of research data management at Loughborough University. It begins with the underpinning research paradigm, discusses the qualitative approach and the case study method and techniques employed to gain evidence, both interviews and documents. What follows is the outline of data collection and analysis. Then, the next section examines constraints such as limitations in methods, data collection and analysis and looks at ethical considerations. Finally, it concludes by analysing the value of research.

Chapter Five analyses and discusses the data of the case study gained from documents as well as interviews to develop understanding of their requirements and opinions about the process.

Chapter Six is the concluding part of this dissertation that examines its key objectives by applying them to

the most relevant findings from the data analysis and discussion. Finally, based on the findings, the last sections make implications for future research and recommendations for Loughborough University and the University Library.

## **CHAPTER 2: Loughborough University**

Loughborough University is located in the UK on a single site campus with 16,106 students across a whole range of disciplines and a total of 3230 FTE staff members (Planning Team, 2012).

The university is one of the leading higher education institutions at a national level and internationally through their accomplishments in teaching and research. They also pride themselves for being the UK's premier university for sports. Their performance has been acknowledged several times and recently the overall university ranking was estimated a 16th place in The Times Good University Guide 2013 and a 14th in the Complete University Guide (Loughborough University, 2012a). Taking into account such factors as student assessment and the quality of research, three disciplines have received top ranking, namely Civil and Building Engineering, Librarianship and Information Management, and Sports Science.

Their overarching mission states "increase knowledge through research" (Loughborough University, 2007) and research plays an increasingly important part on campus. Loughborough University is part of the 1994 Group that brings together research led universities from across the UK, including 19 institutions. Also, they confirmed their research position in the Research Assessment Exercise (RAE) in 2008 that recognised their achievements in research as international (every department) and a world leading (ten departments) (Loughborough University, 2012b).

There is a growing understanding that broadly defined research outputs have a strategic importance to the university enterprise. This responsibility for the research activity is also recognised as all the UK research councils now mandate that research data from publicly funded research be managed and, where appropriate, shared. Hence, the policy of EPSRC has been instrumental to the recent developments at Loughborough University considering research data management (Engineering and Physical Science Research Council, 2012).

According to the current strategic plan, the Research Office plays a significant part in the early phases of research: "Continue to strengthen the research support infrastructure that plays such an important role in the success of the research bidding process and management of external grant income." (Loughborough University, 2007). Described as one of the key facilities, the Library contributes through providing information resources and related services. This focus on information resources is reflected in the key operational documents of the Library. This key document also highlights the role of the Library in the research: "We shall also develop flexible and easily accessible information tools for sharing and developing knowledge through multidisciplinary and internationally-ranging research." (Loughborough University, 2007)

Finally, a close look at the financial position of the university revealed that the research income from

research grants and contracts revealed an increasing tendency according to the Financial Statement 2010-2011 (Finance Office, 2012).

## **2.1 The Loughborough University Library**

Centrally located on campus, the University Library constitutes a single point of reference for the community as it is the only library on campus. The building that hosts the Library measures 7,777 square meters and gives place to 1058 study places and 165 workstations that are spread across three floors (University Library, 2012): the entrance level hosts the Library main enquiry desk, group spaces and study rooms where users are allowed to talk and eat/drink; there is a café and all staff offices. The other two floors invoke more traditional library style – with silent study spaces where users are not able to eat or talk. There are 87 staff members of which 55 FTE (32 full-time and 55 part-time) holding various positions and responsibilities. It should be noted that after a recent reorganisation of the University structure, the Library formed one academic services team and aligned it to the academic schools and departments with targeted support services.

Reflecting the mission and objectives of the University, the Library aims to provide users with access to information resources in a way that is beneficial to them, actively promote its collections, enhance collaboration with support services like the Research Office and across wider campus, and assure the right environment and facilities to the user (Loughborough University Library, 2011). Indeed, the Library provides resources that comprise over 500,000 books, 90 bound serials and approximately 22,000 online journals (University Library, 2012). In addition, once a year the Library undertakes a user study to develop further understanding of user requirements and to enhance the provision of services (Ashby et al., 2011, p. 2). This is also achieved by offering support services, such as dedicated academic librarians who are responsible for communication with schools, its staff and students. They also deliver teaching to PhD and researchers, including workshops and one to one sessions.

In addition to discipline specific services, the Loughborough University Library strategic plan for 2010-2013 highlighted their focus on research support, in particular “support for the varied and changing needs of researchers”. The Library has established a research support group that consists of academic librarians from across disciplines (and schools). The Library Research Support Group is keen on enhancing the role of the Library to the research process and coordinates research related activities through planning, providing information, keeping abreast of developments and maintaining contacts to support research (Research Support, 2011).

Another example of their contribution to the research landscape is through the Institutional Repository. According to a summary outlined by Barwick in 2007, it was set up in 2005 to preserve and showcase the



institutional research according to the Open Access movement. From the outset, this new service was designed to contain full-text documents, including papers, conference contributions, book chapters and electronic theses, and only secondarily, metadata about these documents (2007). A silent aspect of its implementation was the involvement of various research members who participated in the early developments of this service. Thus, it provided insight into different approaches among departments and indicated some of the challenges when dealing with researchers. Most significantly, before this institutional initiative some of the disciplines had already embraced the idea of sharing the research outputs. It also showed a successful involvement of academic librarians in the advocacy for the repository.

## **CHAPTER 3: Literature review**

### **3.1 Introduction**

This chapter offers a review of the literature on the key themes of research data management that crosses various disciplinary fields. Therefore, the publications focus on a selection of arguments, particularly on academic libraries, as well as other stakeholders. Also, where necessary, additional literature is introduced to provide background information and as broad a picture as possible.

In order to cover the abovementioned topics, the literature was retrieved following a specific search strategy. It was conducted on scientific databases such as ACM Digital Library, Emerald and LISTA - Library, Information Science & Technology Abstracts, Web of Science and Google Scholar; the following terms were used and combined by means of search operators: “research data management” OR “data curation” AND librar\*, (management OR curation) AND data AND librar\* in all record fields and without a time limit. Based on the abstracts and relevance to the subject of the research, several journal articles, conference proceedings and reports were identified, and references from these publications helped to obtain additional literature. Nevertheless, this methodology is not without limitations and the researcher does not claim the study to be comprehensive. Searchers were limited to the electronic resources available through Oslo and Tallinn University libraries, and a limited number of databases was included in this methodology. The language was restricted to English.

### **3.2 Definition of research data management**

Researchers have not agreed on a common definition of research data management. Lack of conceptual clarity regards such terms as curation, preservation and stewardship that are underpinning research data management, and no clear distinction exists between these terms. Nevertheless, in this study, the term research data management was employed bearing in mind its limitations.

In one of the early reports on data curation in the UK, Lord & Macdonald used interchangeably data curation and stewardship (2003). Whilst they distinguished between curation, archiving and preservation of data and offered working definitions of these terms, they employed curation to describe the creation, management and long-term care for research data (2003, p. 5), thus using curation as an overarching term (2003, p. 12).

The Digital Curation Centre (DCC) in the UK, which is recognised as “a leader in the area of digital data curation” (Steinhart et al., 2008) also broadly interpreted digital curation:

This view of curation embraces and goes beyond that of enhanced present-day re-use, and of archival responsibility, to embrace stewardship that adds value through the provision of context and linkage: placing emphasis on publishing data in ways that ease re-use and promoting accountability and integration. [...] Digital curation itself is the active

management of data over the life-cycle of scholarly and scientific interest; it is the key to reproducibility and re-use.” (Rusbridge et al., 2005, pp. 2–3)

Reflecting this broad approach their current refined definition states concisely: “Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle.” (Digital Curation Centre, 2012).

Interestingly, the motto of the DCC states “because good research needs good data”. Likewise, Geller (2010), in her discussion of intellectual property around data management, put data management as part of a regular, *good* research practice: “Although some researchers may be resistant to instituting and maintaining good data management practices in their laboratories, such data management, which requires accurate and contemporaneous recording of data, is nothing more than good scientific practice.”

Discussing the role of the academic library in e-research, Thomas (2011) distinguished six types of related activities, namely collaborative technologies, research data management, scholarly communication, computation, visualisation, data collection and analysis. The definition of research data management was limited to different types of data; the author also mentioned data storage and curation activities. (Thomas, 2011, pp. 38–39). Likewise, Lage, Losoff, & Maness defined data curation as a sub discipline of e-science (2011).

Wilson et al. (2010) described research data management as a series of activities along the research life cycle which “involves all the processes that information from research inputs undergoes as it is manipulated and analysed en route to becoming a research output.” (Wilson et al., 2010). They were defined from the perspective of two communities involved in research, those involved in data preservation and curation in repositories and researchers who organise data in databases, files or folders.

In a recent publication Caplan (2012) claimed that the origin of this fuzziness in the usage of terms might be the somewhat unclear term digital curation: “One reason may be a persistent uneasiness with the term digital curation.”

In an attempt to clarify digital curation, Yakel (2007) offered a definition based on major reports and official definitions regarding this topic. Accordingly, she distilled the core concepts of digital curation as:

- (1) Life cycle/continuum management of the materials perhaps even reaching back to the creation of the record keeping system.
- (2) Active involvement over time of both the records creators and potentially digital curators.
- (3) Appraisal and selection of materials.
- (4) Development and provision of access.
- (5) Ensuring preservation (usability and accessibility) of the objects. (Yakel, 2007, p. 338)

Likewise, towards clarity and a shared terminology, Beagrie (2006) provided an overview of digital

curation tracing its meaning back to the origins of ‘curation’ and the purpose which it was supposed to serve – foster relationships between librarians and scientists.

Furthermore, Caplan provided some insight into the conceptual debate claiming that the terms curation and preservation are still being discussed (Caplan, 2012). Her key assumption was that “It seems that in any slugfest between digital curation and data management, data management is bound to win.” (2012, p. 111). Interestingly, literature reveals several examples when these terms were used interchangeably and with different implications, like curation and preservation (Akmon, et al., 2011) preservation and stewardship (Berman, 2008), preservation, management and curation (Hockx-Yu, 2006), and stewardship that involves preservation and curation (ARL/NSF Workshop, 2006). In another contribution, authors used preservation referring directly and indirectly to curational activities (Groenewald & Breytenbach, 2011).

This variety of approaches indicate that different professional communities, including archivists, librarians, computer scientists and curators, have a stake in digital curation, and this is clearly reflected in the recurring terminology. Furthermore, these intricacies have accompanied the debate around digital curation and preservation since the beginning (Macdonald & Lord, 2003, p. 5).

The key element in the abovementioned definitions is data. In fact, research data were recognised as the central part of research work (Heidorn, 2008; Joint, 2007).

As with the management of research data, no shared definition exists for research data; however a Dutch report on the legal status of raw data sought this clarity and provided the two following definitions respectively from humanities and natural sciences:

All data collected in some way or another in the context of scientific/scholarly research. A distinction can be made between primary data (empirical, observed, measured data) and secondary data. Secondary data is data derived from sources created previously (figures published by the authorities, data assembled previously, archived data, texts, etc.).

A datum is an element that has relevance and semantic value...Data is used to describe features of persons, things, actions, etc. taken from reality. (de Cock Buning, Ringnalda, & van der Linden, 2009, p. 5)

Research data consist of a number of types according to context in which they are generated. More generally, they encompass a broad array of scholarly outputs ranging from articles, digitised materials to paper-based notes (Wilson et al., 2010). In some disciplines, like humanities, researchers do not consider themselves to be dealing with data (Wilson et al., 2010). Indeed, Carlson (2012) pointed out that definition varies across communities:

What constitutes data may be interpreted differently by different people at different times. Furthermore, data as a term is often associated with numerical, tabular data by default. Some disciplines, particularly in the Humanities, may not think in terms of working with “data”. (Jake Carlson, 2012)

This author also provided a definition of “data set” as collections of data for a specific project or problem; they may include a variety of data types such as text files, spreadsheets and specific laboratory software files (Jake Carlson, 2012).

Data preservation has been on the agenda since the last two decades (ARL/NSF Workshop, 2006). However, managing research data is not a new aspect of the research landscape. Notably, first data archives were established in 1960s to archive survey data from the social sciences and these were also the first archives of electronic material (Doorn & Tjalsma, 2007). From the beginning, the research archives were supposed to serve three specific aims: data verification, reuse and historical value (2007). Furthermore, early examples of data services in libraries were in Geographical Information Systems (GIS), bioinformatics (Gold, 2007b) and social sciences (Gold, 2007b; Ogburn, 2010).

### **3.3 Drivers of research data management**

#### **3.3.1 Data deluge**

A fundamental shift in data driven research is the ability to generate, compute or collect massive amounts of data that has been emphasised as “Data are becoming the new raw material of business” (“Data, data everywhere,” 2010), which indicates that this topic is now on the agenda.

Recently *Science* dedicated a special issue to data and presented results from a poll in which their collaborators were asked to estimate the size of the largest data set:

We received about 1700 responses, representing input from an international and interdisciplinary group of scientific leaders. About 20% of the respondents regularly use or analyze data sets exceeding 100 gigabytes, and 7% use data sets exceeding 1 terabyte. (Science Staff, 2011).

Rates and volumes of research data are in constant growth. In 2003 Hey & Trefethen (2003) outlined some of the most challenging numbers of the data produced, stored or managed. To illustrate this trend they provided examples from across disciplines stating that for instance in engineering and bioinformatics the order of magnitude of primary data exceeds gigabytes.

This way of conducting research is underpinned by computing technologies combined with advances in communication which are driving widespread developments in science.

The uptake of web technologies, the application of advanced information and communication technologies and the ongoing developments in information technology and computer science have fundamentally changed the way that research is carried out. (Henty, et al., 2008, p. 34).

Reflecting on the use of computers in the research and the new opportunities they can offer, two authors recalled the following expression:

One consistent metaphor in this study likens the computer to a moveable and adjustable lens that allows scholars to

view their subjects more closely, more distantly, or from a different angle than would be possible without it. (Williford & Henry, 2012, p. 21)

Developments in the scholarly practice are variously described as a new age of science (Goldenberg-Hart, 2004; Soehner et al., 2010), the new science paradigm (Paterson, et al., 2007) and a new collaborative science (Hey & Hey, 2006; Hey & Trefethen, 2003; Lord & Macdonald, 2003) that is data driven and employs information and communication technology (ICT) to organise and manage research, and conduct collaborative projects. Along with various perspectives these developments have been called variously e-research, e-science, cyberinfrastructure, and digital humanities.

According to Lynch, technological opportunities are not limited to a number of research fields but rather cut across disciplines (Lynch & Carleton, 2009); still scientific research reveals some patterns across disciplinary areas that align to these changes. On the one hand, ‘big science’, a term that is used to indicate large scale collaborative and cross-institutional research projects that use computational power to generate massive and potentially unlimited amounts of data from instruments, measurements and during experiments; it comprises such disciplines as climate, astronomy, genomics (S. Carlson, 2006) high energy physics, meteorology (Swan & Brown, 2008), engineering and bioinformatics (Hey & Trefethen, 2003). On the other hand, ‘small science’, whilst apparently less demanding, these research projects increasingly produce and access big amount of data (Humphrey, et al., 2000). According to Heidorn, research conducted at a smaller scale attracts fewer resources and less attention (2008). He argued that the majority of research work is done in relatively small projects, which are however not synonymous with small science (Heidorn, 2008). A third interdisciplinary and collaborative approach brings these two ways of conducting research together. In a study on computationally intensive research, Williford & Henry, put in evidence the existence of what they called one culture:

In assessing the project teams’ work, we have come to understand that the one culture of e-research – encompassing what have been called the e-sciences, as well as the digital humanities – involves not a choice between the scientific and humanistic visions of the world, but an imperative that people and organizations fully embrace both. (Williford & Henry, 2012, p. 7)

The volume of data opened up new opportunities of conducting research and in the same time raised some challenges to the way research is carried out. It not only increased the quantity but also made a qualitative difference (“Data, data everywhere,” 2010).

To illustrate the impact of e-research on researchers Zhao (2009) highlighted a broad spectrum of possibilities, ranging from collaboration, through sharing resources and data over the network, higher levels of data processing and managing, advanced understanding of research, the possibility to use all potential resources, more efficient research in terms of cost and time to a facilitated access for students to

research. Another area most commonly known where some new, sophisticated experiments were made available is particle physics and in particular the Large Hadron Collider project (Hey & Trefethen, 2003; Lord & Macdonald, 2003).

However, computationally enabled datasets can be very challenging because of their sheer volume and complexity (Hayes, Harroun, & Temple, 2009). Furthermore, preservation of these data is by far uncertain:

Leaving digitally based information to languish in personal electronic filing drawers amid a jumble of unrelated information and with no plans for its survival guarantees its disappearance. Unlike the upkeep of our academic buildings, deferred maintenance is not an acceptable strategy for preserving data. (Ogburn, 2010, p. 242)

Ogburn also concisely examined major challenges regarding digital information and research data, and related these issues to the intrinsic features of knowledge in digital form that can be “unreadable, corrupted, erased, or otherwise impossible to recover and use” (Ogburn, 2010, p. 242).

Research projects that increasingly rely on data look at the quality of these data to a greater degree. Indeed, with a numerical growth, there is a perceived recognition of the value of research data (Beagrie, 2006). Williford & Henry acknowledged:

The quality, quantity, and utility of data is unquestionably the most complex determining aspect of these projects [...] it is not just the specificity of the question or the maturity of a tool that determines what computationally intensive research might achieve, but also the state of the raw material from which it is produced. (Williford & Henry, 2012, p. 14)

### **3.3.2 Research practice**

Whilst data intensive research is already part of the history (Lord & Macdonald, 2003, p. 9), overwhelming volumes of data are still a challenge for researchers. Dealing with data requires additional resources, such as time and skills. “[Researchers] will need the skills and technologies both of computer scientists and of the library community to manage, search and curate these new data resources.” (Hey & Hey, 2006). Cragin et al. observed that researchers seem not prepared adequately to conduct the new research: “Many scholars are unaware of the coming changes in the sociology of science and do not have the required skill sets to address the requirements in their new proposals (Cragin, et al., 2010).

Hayes et al. (2009) notably pointed out that there are no requirements or lack of funding for managing data. They also observed that that the majority of data are managed by graduate students and post doctoral fellow holding temporary positions within universities. In this regard, Williford & Henry appraised the role of individuals who manage data on behalf of regular researchers (2012).

There is a consistent body of literature exploring the needs of researchers (Lage et al., 2011, p. 915). From an analysis of studies that aimed to assess requirements and practices of researchers, it is evident that scholarly disciplines have varied and diverse data practices (Beagrie, Beagrie, & Rowlands, 2009; Jake

Carlson, 2012; HATII & University of Glasgow, 2009; Wilson et al., 2010).

Mainly, problems regard researchers whose practice may place data at risk (Ogburn, 2010); Joint argued that they are dealing with data following their own rules (Joint, 2007).

Studies concerning assessment of needs provided evidence on similarities and common features. Henty et al. (2008) undertook a study that offered an insight into researchers' needs across three Australian universities showing some similarities across the three institutions. These were coexistence of digital and paper-based data, increasing amount of digital data, variety of data types and related software, diverse storage solutions, lack of data management plans, a variety of researchers' perception on the responsibility for data, approach to data sharing, and variety of means to accessing and re-using data. In addition, a study carried out by HATII and the university of Glasgow identified common data issues:

- “Poor naming and filing systems so retrieval is a challenge;
- Lack of storage space meaning researchers revert to using external hard drives and laptops;
- No active transfer of data on staff retirement / departure meaning legacy material is lost, mismanaged or remains on the server unused;
- Limited data archiving facilities, so researchers often have to maintain their research outputs;
- Growing requirements e.g. from publishers and RCs [Research Councils] that researcher feel ill-equipped to meet” (HATII & University of Glasgow, 2009, p. 9)

Reflecting on data practice from the point of view of Information Science, these studies revealed another common theme, namely the need for timely training and advice (Delserrone, 2008; Henty et al., 2008; Thomas, 2011; Wilson et al., 2010). This is supported by the findings from other studies (HATII & University of Glasgow, 2009; Peters & Dryden, 2011) that found that researchers need assistance and support to adopt best practice through data lifecycle and address data management plans. In addition, there were common features when developing infrastructure “The need for data documentation (metadata), training and support, secure storage, and linking data to publications are common across disciplines.” (HATII & University of Glasgow, 2009).

There is a shared understanding that the management of research data is essential for advancement of science (ARL/NSF Workshop, 2006) and is required because of underlying technology, social factors and organisational risk (Beagrie, 2006).

### **3.3.3 Research councils**

The increase in importance of data management directly influenced funding agencies that appreciated the value of research data as their return on investment and as the most valuable element of research practice (Caplan, 2012).

In fact the main rationale behind research data management is a growing demand to provide a broad



access to publicly funded research. Such international body as Organisation for Economic Co-operation and Development (OECD) acknowledged in their seminal guidelines for access to research data from public funding that access to these data is fundamental for scientific research (OECD, 2007). Similarly, national institutions worldwide were committed to address the need for access to and enhanced visibility of publicly funded research (Paterson et al., 2007).

Funding agencies are increasingly requiring publicly funded research to be shared in open and largely available repositories (Jake Carlson, 2012). In the UK, research councils now mandate the deposit and sharing of data and this is according to the Research Councils UK (RCUK) principle on data sharing:

Making research data available to users is a core part of the Research Councils' remit and is undertaken in a variety of ways. We are committed to transparency and to a coherent approach across the research base. These RCUK common principles on data policy provide an overarching framework for individual Research Council policies on data policy. (Research Councils UK, 2012)

Reflecting these principles and in light of data protection, the key UK research council, Engineering and Physical Sciences Research Council (EPSRC), has set up a policy on research data management that states:

firstly, that publicly funded research data should generally be made as widely and freely available as possible in a timely and responsible manner; and, secondly, that the research process should not be damaged by the inappropriate release of such data. (Engineering and Physical Science Research Council, 2012)

In fact, this policy frames how research grants are distributed, which has an influence on researchers, research institutions and the general public. First of all, researchers are required to manage and share their data, to inform about the research data related to a publication, including digital and paper formats. Secondly, institutions are responsible for raising awareness and informing the research community; thus, they need to prepare adequate framework for research data holdings, namely access, management, preservation and lifelong curation of data, and publish metadata for the research data that respond to the four essential questions what, why, when and how. Finally, funding bodies are responsible for financial support to sustain the research data. However, it is not clear how they would estimate the cost of such service and to what extent would they sustain it (Engineering and Physical Science Research Council, 2011).

This trend is a logical consequence of the Open Access movement and is part of a wider discourse on scholarly communication (Accart, 2011). Based on the principle of free and unrestricted online access to human knowledge and in light of the Berlin Declaration, Open Access considered raw data along with other scholarly outputs as part of global scientific knowledge (Berlin Declaration, 2003).

### **3.3.4 Academic libraries**

Central to academic libraries is managing, providing access to and preserving information and they have

been recognised as a trusted partner in the information and knowledge cycle. In this regard Lynch mentioned the following core attributes of research libraries: collecting, organizing, providing access, and archiving (Lynch & Carleton, 2009, pp. 240–242). Libraries have also contributed significantly to provide intellectual access, metadata and classification schemas (Lankes, et al., 2008).

Furthermore, Osswald argued that traditional library competences and digital library (technical information) services, were substantial to enhance the quality, add value and enrich e-science projects and provide adequate services (Osswald, 2008). In a similar manner, Joint observed that professionals may enhance the research activity:

In particular there is a danger that scientists themselves will apply the information management techniques of the new science to their own activities inappropriately. [...] Scientists who are very much “within” their own data have little overview and sense of perspective as a result of this inevitable tunnel vision. (Joint, 2007)

Therefore, it was broadly recognised that libraries play a prominent role in managing research data. During “E-Research and Supporting Cyberinfrastructure” forum organised jointly by the Association of Research Libraries (ARL) and the Coalition of Networked Information (CNI), Lynch (as cited in Goldenberg-Hart, 2004) mentioned that libraries have a central role in addressing e-research; together with IT, libraries were described as key players in the e-research (Association of Research Libraries, 2004).

The initial debate also evidenced that a lack of adequate services would put at risk the entire institution (Lynch as cited in Goldenberg-Hart, 2004); thus, if libraries approach timely this issue, they can redefine their role in support of research and gain a more prominent position in the research landscape (Association of Research Libraries, 2004). Hey & Hey (2006) also stressed this responsibility of libraries: “the e-Science revolution will put libraries and repositories center stage in the development of the next generation research infrastructure”. Accordingly, Swan & Brown observed: “The role of the library in data-intensive research is important and a strategic repositioning of the library with respect to research support is now appropriate.” (Swan & Brown, 2008).

Furthermore, overall responsibilities of academic libraries with regards to research data were described as maintenance and curation, as mentioned by Hey & Trefethen (2003); Swan & Brown acknowledged the need to archive and preserve research data. Brandt clearly pointed out that librarians need to get involved in data management and related activities: it has become obvious that there is an urgent need for the knowledge that librarians have: the ability to collect, organize, describe, curate, archive, and disseminate data and information. (Brandt, 2007)

These arguments were setting the stage of debate around the role of libraries in the research landscape. However, questions about roles, responsibilities and competences of libraries and librarians persist:

As newly emerging e-research practices increasingly harness resources and services offered by academic libraries, these libraries are compelled to rethink their workforce planning and service delivery models to incorporate these new practices. (Thomas, 2011).

A few authors pointed out that libraries need to extend and adequately adapt services they offer (Genoni, Merrick, & Willson, 2006; Kallenborn & Becker, 2009; Ogburn, 2010). Kallenborn & Becker observed: “the conveyance of expert knowledge can no longer be confined to OPAC statements or the presentation of search strategies for specialist databases” (2009). Put in other words, traditional services focused on publications in form of articles and monographs are not enough to support current research needs, and there is a need for an innovative approach. However, Heidorn observed that there is still much to clarify, especially regarding practitioners, curation responsibilities and economic frameworks:

librarians know relatively little about current data management practices of scholars. Institutions have not yet established who will conduct data curation work. Funding models are still being developed and the economies of scale must be examined. (Heidorn, 2011, p. 665)

In the debate around the role of academic libraries in the research practice and research data management, a few authors highlighted the prominent role that librarians, variously called subject, reference and liaison, who directly engaged with researchers, play in the research data lifecycle (Auckland, 2012; Newton, Miller, & Bracke, 2010; Witt, et al., 2009). Auckland (2012) carried out a study on the changing needs of researchers in the context of library offerings and evidenced their good subject/domain knowledge and the ability to understand researcher or research project needs.

On the whole, reference librarians were recognised as capable of addressing key requirements of the research community:

The ability of reference librarians to work both within and across disciplines, to develop trusted relationships with faculty based on an understanding of their individual needs, and to cross administrative boundaries and bring different constituencies together are key elements in addressing the challenges described in working with data. (Jake Carlson, 2012).

More specifically, this author observed that librarians can further engage with researchers as data interviewers in order to investigate the research process and develop understanding of data practices and workflows (Jake Carlson, 2012).

Apparently, authors do not agree who specifically should hold this responsibility of engaging with researchers in order to develop further knowledge of the research process and research data. In some publications, authors discussed professionals actively engaged in the research practice but using alternative terms such as data archivist (Humphrey et al., 2000), data specialist (Lewis, 2008; Newton et al., 2010), data liaison (Gabridge, 2009) and an envisioned role of a cyberinfrastructure facilitator (Lankes et

al., 2008). Even more, Choudhury added data scientist and data humanist to this poll of data professionals, emphasising that these may be a future for subject librarians (2008).

In more general terms, Hey & Hey (2006) mentioned collaboration between researchers and librarians highlighting collaborative and communicational skills; Candela et al. (2009) described the role of librarians as working closely with researchers and stated that this would require specific domain knowledge. In fact, what the abovementioned roles share with these more general roles, is domain knowledge and data management expertise. Less frequently authors mentioned technical skills (Alvaro, et al., 2011; Jacob Carlson, et al., 2011).

A complementary observation was made by Brophy (2007) who supported this requirement to understand the research practice and subject knowledge; based on the constructivist approach, his arguments were in favour of a shared meaning and communication between researchers and librarians: The issue is not just about terminology but more critically about concepts and interpretation. It takes us beyond data, beyond information, and beyond knowledge towards shared meaning. (Brophy, 2007)

The aforementioned roles may seem new as some of these examples illustrated. However, upon reflection these roles imply contrasting opinions. Hswe & Holt claimed that many of data related roles, responsibilities and activities had been familiar to libraries and need not be developed anew: “there is arguably much territory already charted here—to the extent that some of it may need only to be remapped toward either specific or generalizable uses.” (Hswe & Holt, 2011, p. 12). Furthermore, they observed that the aforementioned argument advanced by Gabridge (and other authors) about liaison roles regards distinct roles within libraries:

Implicit in Gabridge’s foregoing appeal is the depth and range of librarian expertise that cuts across boundaries of practice and skill sets: subject specialists, metadata librarians, institutional repository coordinators, data curators, systems/IT librarians, copyright specialists, collection managers, and acquisition librarians (for advisement on data sharing and collection policies). (Hswe & Holt, 2011, p. 13)

Alvaro et al. drew similar conclusions in their study on library involvement in e-science that was based on literature review and library employment advertisements: “These titles reveal the different focuses of the positions such as data, metadata, liaison responsibilities, scholarly communication, and specific scientific subject knowledge. If the field were defined, one would expect some consistency in the job titles, but what was found was not two different terms being exchanged in and out, but titles referring to different positions that would coincide with e-science in different ways.” (Alvaro et al., 2011) In conclusion, they identified three different roles related to e-science: data librarian, subject librarian and e-science librarian.

Still a slightly contrasting and complementary argument came from Beagrie who recognised that digital curation requires skills of different stakeholders and roles crossing professional boundaries (Beagrie,

2006).

In a recent study on the changing needs of researchers in the context of library offerings Auckland mapped some of the researchers' needs against skills and services which are required to meet those needs (Auckland, 2012). By exploring the current situation, the author highlighted some of the skills and knowledge regarded as increasingly important but which are currently lacking. Particularly relevant to the topic of research data management were preservation of research outputs, data management and curation, funders requirements, data manipulation tools, data mining, metadata, preservation of project records, sources of research funding to assist researchers to identify potential funders, metadata schema, and discipline/subject standards and practices (Auckland, 2012, p. 43).

As previously mentioned, some authors (Gold, 2007a; Ogburn, 2010) identified a number of data services familiar to libraries; Gold identified several reference librarians who have already incorporated data services into their work; Genoni et al. noted that some institutional repositories may contain informal categories of content such as unpublished research material including data (Genoni et al., 2006). It has been observed that these datasets can constitute special library collections (Choudhury, 2008). In addition, Angevaere claimed that libraries might take these on new types of content, in order to regain their unique position, given that they had lost responsibility in the curation of journals (Angevaere, 2009).

A scope for library intervention lays in a wider involvement of librarians throughout data lifecycle and their collaboration with researchers. According to Gold (2007a) librarians can provide assistance as early in the research practice as data discovery, selection, acquisition, and licensing and they have the opportunity to support metadata documentation, share best practices and standards.

Recent developments indicate a shift from archiving, which apparently denotes a passive responsibility (Beagrie, 2006), to active involvement in the research activity of those who can be potentially involved in data curation and researchers. Doorn & Tjalsma evidenced these significant aspects:

two developments—placing the responsibility for preserving research data for the long term in the hands of the researchers themselves and the need to provide data with lifelong care—have led to new ideas on the data infrastructure (Doorn & Tjalsma, 2007).

For the above reason to provide curation over a long period, the key conclusion that Gold drew was about the involvement of librarians upstream, in the early stages of research work (Gold, 2007a). In her opinion other potential actions that librarians can take to integrate data services into their current remit, were through developing and marketing data consultancy and referral services (Gold, 2007a).

Furthermore, librarians have a recognised capacity in training and assistance. Reflecting upon the role of academic libraries in learning and teaching, Brophy argued that these activities are also their key role: “at

the most fundamental level, academic libraries are in the business of human learning rather than in the information business.” (Brophy, 2007). Current provision of information literacy can facilitate development of data information literacy (Ogburn, 2010). Carlson et al. observed:

Most academic libraries already offer information literacy courses and programs as a part of their services. Extending these information literacy efforts to include programs on data management and curation may provide a logical entry point into increasing libraries’ role in supporting e-research. (Jacob Carlson et al., 2011, p. 630)

Besides, specific contribution that librarians can make to wider support research practice is through data repositories. Libraries can offer new repositories or extend their repositories to include a broad variety of scholarly outputs, such as raw data. Indeed, Berman recognised that critical research data require to be preserved over a long term and institutions hold the responsibility for community repositories (Berman, 2008). Discussing the role of institutional repositories, Choudhury claimed that data curation was one of the “repository-related services” and recognised the role of repositories in data intensive research. The author suggested that institutional repositories would be instrumental to support new outputs and strategic towards a preservation of these outputs. (Choudhury, 2008).

According to Paterson et al. (2007) there have been two waves of developments in the research landscape. During the first phase institutional repositories were created to collect and share scholarly publications, and the second took into consideration also other research outputs such as raw data. However, this is in contradiction with what other authors stated regarding the repositories of electronic materials (Doorn & Tjalsma, 2007). The next step is linking the two, namely providing access to and sharing of final publications along with original datasets. For instance, the university of Southampton was planning to build a research data platform and provide an environment in which publications and original data will be linked, and data would be citeable (Simpson & Hey, 2006).

### **3.4 Implementation of research data management**

Whilst a number of institutions are still considering research data management and quite a few libraries are negotiating their role in these developments, the following initiatives were particularly worth mentioning.

In an attempt to assess current data practices and define the role of librarians in research data management, Peters & Dryden (2011) found out that the library was capable of adequately supporting the new process as, rather than a technical infrastructure, researchers were seeking assistance and support with their data practice. Their investigation had also broader effects regarding researchers’ perception of the library:

Many of the faculty who fall within this latter category [do not have regular contacts with the library] have been surprised to find that the library can offer research support services other than traditional library instruction. (Peters & Dryden, 2011, p. 398)

Furthermore, their planning for future activities included further assessment of researcher needs focused on graduates and post-doctorates, who were judged as the key responsible for data activities, to set up a working group to deliver a consistent message and develop a host of services under the responsibility of the library (Peters & Dryden, 2011).

Some libraries developed real life examples of librarians partnering directly with researchers, such as the data curation profiles, which were developed at Purdue University and the personas, or researchers' profiles, as described in (Lage et al., 2011).

Thomas (2011) explored how the library prepared themselves to support e-research, including the library's roles and responsibilities in developing services that support e-research (data librarian, research support librarian). Similarly, in a case study of the University of Minnesota as described by Delseron (2008), the library hired science librarian cohort, three science librarians to investigate interdisciplinary research collaborations occurring within and between various research arenas.

In addition, one author highlights the need to provide local research data services and to define what should be provided locally and which data related activities would be done at a national or international level (Lewis, 2008). Local needs and fitness for purpose were also among the principles of a managing research data pilot project at the University of the West of England (Fowler, 2012).

When setting up a service such as research data the library needs to learn from the experience of institutional repositories if the service aims to support research (Jake Carlson, 2012; Martinez-Urbe & Macdonald, 2009; Newton et al., 2010), and work practice of professionals involved in implementing it. Institutional repositories were not populated according to the initial expectations and some observed that researchers were not involved in the planning for and development of these repositories (Salo, 2007). Thus, Martinez-Urbe & Macdonald argued that direct engagement of researchers was the key requirement when building systems to organise, manage and store research data (2009). In fact, the use of such infrastructure is dependent on their perceived necessity, thus depends on the interest of researchers in handing in their data to an outside party (Lage et al., 2011). Accordingly, researchers were placed at the heart of developments (HATII & University of Glasgow, 2009; Martinez-Urbe & Macdonald, 2009; Wilson et al., 2010).

Furthermore, Van de Sompel et al. aptly reminded "Like any technology, success will depend not only on technical soundness but on the willingness of the participants in the system—publishers, scholars, academic institutions, funding institutions, and others—to adopt new tools and develop new organizational models on top of them." (2004). Indeed, some authors described concerns regarding librarians who were not advocating or were not fully supportive of institutional repository. Therefore, a

successful implementation is strongly related to support and advocacy of involved librarians. (Dorner & Revell, 2011)

The literature broadly described that data support needs to be provided in a collaborative effort (Goldenberg-Hart, 2004; Newton et al., 2010) and this would enable a successful undertaking (Hey & Hey, 2006). Hey & Trefethen observed that “a solution to these problems [data deluge] is much more than just a technical challenge: all parts of the community from digital librarians and scientists to computer scientists and IT companies need to be involved” (Hey & Trefethen, 2003, p. 14). Authors also pointed out the unique position of libraries to make connections between different members of the research community (Choudhury, 2008) and build partnership (ARL/NSF Workshop, 2006).

However, Wilson et al. noted “recognising that research data management belongs to no single part of the institution but rather involves a partnership of researchers, their departments, IT, library, and research support services” (2010). He further observed that there are no obvious choices as to which cohort or support service would take the prime responsibility for research data management:

It matters less which part of the organisation is taking a lead on these activities than whether the relevant providers are engaged in the undertaking and have a reasonably clear sense of not only their current service provision and strengths but also their gaps and weaknesses. (Wilson et al., 2010)

Of particular importance is the partnership between librarians, or those involved in the research process, and researchers themselves. However, this poses some questions, as to the practicalities of user engagement, such as choosing the research community or group (McKay, 2010). Lage et al. observed that strategically this partnership would involve researchers receptive to the library’s involvement (Lage et al., 2011, p. 916). Bracke further discussed some solutions for engaging with researchers and invited to experiment with researchers who are interested in sharing data and whose data are in formats that do not require complicated technical solutions (2011). In fact, one of the key elements of a successful implementation of new services are institutional collaboration and inter-institutional partnership (Mandel as cited in Goldenberg-Hart, 2004).



## CHAPTER 4: Methodology

### 4.1 Research paradigm and approach

This research is characterised by a set of assumptions that drive the choices of this study. First, social world consists of individual worlds rather than universal reality, in which the point of view of individuals and their context play a prominent role; thus is influenced by the perspective of constructionists as presented in Denscombe (2009, p. 113). Second, knowledge is created in interaction and by means of transactional processes in which all elements interact with each other, and this can be described as “simultaneous shaping” (Pickard, 2007, p. 12). Finally, methodological approach is based on empathetic interaction between parties involved and context plays once again a fundamental factor. “Context is something you swim in like a fish” (Dervin 1997, 32) (as cited in Pickard, 2007, p. 12).

These implications are fundamental to understand appropriately the conceptualisation of this study, as it is concerned with developing knowledge about individuals. Accordingly, qualitative approach was chosen to enable social construction of reality as outlined in (Gorman, et al., 2005). Furthermore, the research question, specific aims and, particularly, the interview technique provided a solid ground for an in-depth investigation of individuals.

Although researchers hesitate to agree on a definition of qualitative inquiry, it can be determined from its main attributes as outlined in the literature. The following seemed the key aims: develop a deep understanding, provide a rich, descriptive picture and contextual information (Trochim, 2006), lack of detailed procedures at the outset of research (Denscombe, 2009), and draw data from the context (Gorman et al., 2005).

Again, as with metatheoretical statements acknowledged above, the key assumption in the qualitative approach is that the understanding of a phenomenon under investigation can be gained only from its own perspective. Indeed, according to Gorman et al. “the meaning of events, occurrences and interactions can be understood only through the eyes of actual participants in specific situations” (2005, p. 3). Essentially, in this methodological approach participants are the source of valuable meaning. However, their contribution and the overall understanding are highly dependent on the researcher, who is regarded as the research instrument. “Human lives and their interpersonal relationships create complexities that need to be understood and the researcher acting as the research instrument allows for understanding and depicting these complexities” (Pickard, 2007, p. 14). Discussing the choice of research methods Gillham aptly noted “All raw data require interpretation” (Gillham, 2005, p. 8). In fact, this further highlights the importance of mutual influence of those taking part in the research and its role in qualitative research.

Final salient aspect of this approach is its emergent design which allows for flexibility and to adapt

appropriate techniques along the research process. Furthermore, it gives the participants an opportunity to interact with the researcher and to negotiate the outcomes of the inquiry, which adds additional insight into the case being investigated (Pickard, 2007, p. 17).

## **4.2 Research method**

Case study method has been applied to this investigation because it has been regarded as the most suitable for the purpose of the study to address the research question and specific aims, to get a better understanding and in-depth investigation of the target communities, and because it allows for flexibility in the design (Pickard, 2007, pp. 85–93). This choice will be further discussed in *Constraints* section, as right methods and adequate research question are the foundation of this study.

Likewise the definition of qualitative research, case study research has no common definition and difficulties in reaching a shared understanding were outlined in the literature (Gillham, 2005, p. 167; Pickard, 2007, p. 85). Gerring summarised these issues offering a concise definition: case study research is an intensive study of a phenomenon within set boundaries that aims to transfer the findings to other units: “the intensive study of a single case where the purpose of that study is – at least in part – to shed light on a larger class of cases (a population)” (Gerring, 2007, p. 20). Furthermore, this unit of analysis that might be single or multiple does not automatically translate to other cases: “the unit(s) under special focus is not perfectly representative of the population, or is at least questionable” (Gerring, 2007, p. 20). To put it in other words, this method of enquiry is used to examine an instance of a contemporary phenomenon within set boundaries of its context in order to get a holistic understanding of the case. This is obtained through rich description and by providing contextual information.

Accordingly, this study focused on research data management, which is currently on the agenda in some academic libraries; it analysed and discussed a single case in a defined setting, namely Loughborough University and in particular a group of stakeholders that are involved in the implementation of research data management at the university. To guarantee an adequate depth of investigation and, where applicable, to derive knowledge from the case under enquiry (Gorman et al. 2005 chapter 4), this study provided an essential and rich description of the bounded system (Denscombe, 2003, p. 271). Finally, participants were given an opportunity to negotiate the outcomes of the enquiry through a verification process as the researcher contacted the key informant to validate the data analysis.

## **4.3 Sample and sampling methods**

The main purpose of sampling is to get a balanced, representative picture of the unit under investigation (Gillham, 2005). However, Denscombe pointed out that in the social sciences sampling involves a limited number of units for their particular qualities and whose choice depends on a given purpose (Denscombe,

2009, pp. 181–195). Furthermore, in a research that deals with human subjects researcher builds a relationship with the study participants that is based on trust, as highlighted in Pickard (2007). “Gaining entry does not only include the formal aspects of signing off and gaining permission, it also includes establishing trust and building up a rapport with all of the stakeholders: participants, informants and gatekeepers” (Pickard, 2007, p. 88).

A two stage sampling strategy was employed. In the initial plan, the choice of participants was made according to snowball sampling. Working with the key informant, the profiles of potentially information-rich individuals were identified, and some of them were approached personally to establish the first contact during the internship period of the researcher at the Loughborough University Library. Subsequently, the same individuals were invited to participate in the study. In addition, in the second stage, using purposive sampling to gather multiple perspectives, other individuals were identified and similarly invited to participate in the study.

The population of the study included all Loughborough University staff involved in the implementation of research data management, namely the Library, and especially the Library Research Support Group, the Research Office, High Performance Computing, Intellectual Property Office and IT Services. The researchers were excluded from the investigation. This was due to different reasons, the most important being that the literature provided a thorough account of requirements and opinions of researchers regarding their research data practice. Besides, the researcher conducted a pilot study on research data management at the Loughborough University Library in which engaged with a small sample of the research community. More significantly, they form a population which by far exceeds the scope and resources of this study.

The specific unit of enquiry was selected for several reasons, the main one being researchers’ familiarity with the University and the Library. It was crucial as the rationale for this study was to gain an understanding of opinions. Given an online methodology employed for the most part of the fieldwork, gaining entry into a community without personal contacts would be unlikely. Similarly, Tellis highlights the role of ‘learning’ outcomes and time constrains, factors which guide the researcher in the selection of the case for investigation (Tellis, 1997). Equally relevant was the fact that Loughborough University is research led and, as a consequence, the Library supports this mission.

Briefly, the case study has been selected based on the following factors:

- Access to key informants willing to participate and who can provide rich evidence
- Research led academic library
- Research data management is on the agenda

In total, out of nine individuals contacted personally and through email invitations eight members of

Loughborough University participated in the interviews. They were on different positions and from different departments; the majority of them was affiliated with the Library (4), and the other four participants were from the Research Office (2), from the Research Computing (1) and IT Services (1).

**Table 1 Interviewees**

Interviewee	Affiliation	Interview type
A	Library	Email
B	Research Office	online interview
C	Library	online interview
D	Library	online interview
E	Research Computing	Email
F	Library	Email
G	IT Services	online interview
H	Research Office	Email

#### 4.4 Data collection techniques

At the outset of the project, after having identified the key informant, the researcher approached individuals regarded as information-rich for a preliminary conversation on the topic (Pickard, 2007, p. 89); some of them were subsequently recruited for the interviews by email (see Appendix One). In a similar manner, other potential participants were invited to contribute to this study.

The fieldwork was carried out in part on site at Loughborough University, where the preliminary conversations took place and, for the remaining part, in the virtual environment. An online investigation implies considered differences as compared with traditional ways of conducting research (Pickard, 2007, p. 84) and these implications are discussed in *Constraints* section.

##### 4.4.1 Interviews

Peräkylä & Ruusuvuori (2011) observed that this technique is widely used in social sciences research and other disciplines. In particular, they are regarded to be the most important source of information in case

studies (Tellis, 1997). "Most qualitative research probably is based on interviews. There are good reasons for this. By using interviews, the researcher can reach areas of reality that would otherwise remain inaccessible such as people's subjective experiences and attitudes." (eVALUED, 2006)

Recognised as a valid instrument for gaining insight into the views of individuals (Kvale, 1996) and in depth investigation, interviews are also used to validate, "qualify and interpret", other sources (Gillham, 2005, p. 167).

The abovementioned reasons guided the choice of interview as the key data source in this investigation. In fact, this study sought to reveal opinions of individuals, collect in depth information and depict the context thoroughly. Interviews were as well regarded as a relevant source of data to corroborate the findings from the secondary source, namely documents.

Kvale (1996), defined research interview as a conversation between partners interested in a topic: "The research interview is an interpersonal situation, a conversation between two partners about a theme of mutual interest" (p. 125) and "knowledge is created inter the points of view of the interviewer and the interviewee" (p. 124). This approach that Kvale called creating knowledge implies that those contributing to the interview have an influence on it and on the data that emerge from it. For this specific reason in the interview situation: "The interviewer is the research instrument" (Gillham, 2005, p. 7), similarly as in the qualitative approach.

Among the ways of conducting interview, semi-structure interview was selected both for its flexibility and structure (Kvale, 1996, p. 124). Kvale (1996) suggested to design in advance an interview guide to prepare the same series of questions in order to cover common elements but, where needed, giving enough room for additional questions, thus allowing for spontaneous flow of conversation. In practice, during the fieldwork, interview guide and guiding prompts (see Appendix Six) as suggested in (Gillham, 2005) were used to ensure that some topics were covered with all the participants and to help rephrase if the questions were not clear; guiding prompts are also helpful to anticipate the data analysis (Gillham, 2005).

In the original research design, it was envisaged that all the interviews would be held by means of online videoconference. However, research process is not straightforward and, as Pickard pointed out, at the outset of qualitative research it is almost impossible to select the techniques to be used for data collection (Pickard, 2007, p. 89). After the first contacts, some of the potential participants expressed doubts concerning this type of interview, as they needed to arrange it. As an alternative, they were offered an email interview, which would allow for convenient time to respond and would not require additional arrangements. On the whole, it was regarded as less intrusive and the participants could decide more freely when to respond to the research questions. Besides, the researcher was offered additional flexibility in

reading and interpreting the responses.

As a result, four interviews were conducted using a videoconference and the same number of email interviews was held. Furthermore, such flexibility should have been guaranteed from the outset of the research giving all the participants the possibility to choose the most suitable data gathering technique and channel.

Concerning the ways of conducting interviews distant methods were chosen, such as the aforementioned online interviews, that responded to the practical matters, namely time, resources and access to relevant data and interviewees (Gillham, 2005).

Authors distinguished between asynchronous and synchronous ways of communication comparing them to the traditional means. Synchronous online interviews share similarities with face-to-face interviews, as they are spontaneous, use oral account and provide extra-linguistic information, like gestures and facial expressions. James & Busher presented a comprehensive comparison of several features of face-to-face and online interviews (synchronous and asynchronous) covering cost, access, temporal dimension, nature and speed of response, time, venue and participation, quality of data, identity and confidentiality (James & Busher, 2009 Table 1.3). Whilst the authors focused mainly on asynchronous means, some aspects of synchronous interviewing can be derived from the features of face-to-face approach. They are: happening in real-time, spontaneous, constrained by time, rich in visual and verbal cues, thus improve mutual understanding (2009 Table 1.3).

Regarding online interviewing by means of videoconference, Gillham (2005) discussed some significant aspects of video interview, which shares similarities with online interview. Video recording is a complete account of the interview and can be used as a form of reference to validate findings, or an 'audit trail'. Also, it consists of both voice and video and guarantees similar qualities as traditional means of communication.

Email interview is regarded a valid way of getting rich data (Gibson, 2010). Discussing its strengths and weaknesses Bampton & Cowton (2002) highlighted two major displacements of time and space, as the interaction is mediated by a computer, is screen-based and takes place in different times. For the convenience of busy subjects, e-interviews do not necessitate "mutually convenient time" (Bampton & Cowton, 2002). In addition, one fundamental characteristic of email interview is an iterative way of exchanging emails, which characterises any good interview.

Accordingly, this enquiry used one of the available audio-visual online tools for interviewing, Skype, to enable a good mutual understanding. Alternatively, those participants with a busy agenda received an email with research questions in the body of the message (see Appendix 5). They not only responded to

the questions but also offered to answer additional questions, even if follow up communication happened only in one case, when it was regarded as necessary.

#### **4.4.2 Document analysis**

In this study not only primary sources of information have been selected, but also secondary sources, such as documents relevant to Loughborough University and the University Library. They provided an insight into current services and planned activities, were used as a background for the interview material and to validate these empirical findings. This is in line with the aim of this study to examine current services and planned activities.

Initially, a review of institutional documents was conducted selecting those pertinent to the study. They included administrative records gained during work on site, thus mainly focused on the Library, as well as reports available through the University web site. Further analysis included policies, strategies and other institutional documents relevant to research data management at Loughborough University. Where necessary, policies of research funding bodies were analysed to obtain contextual information.

#### **4.5 Data analysis**

Qualitative data analysis was used. The major source of data were interviews, other data sources being documents, therefore qualitative analysis was regarded as appropriate.

The data analysis took place along the data collection to allow categories to emerge from the data (Pickard, 2007, pp. 155–163). Interviews were transcribed, and this transcription “translations” (Gillham, 2005), was done as soon as possible after the interview for ease of interpretation and to be able to learn from one interview before the next one starts. What followed was editing of the transcripts and an edited account of interview material was sent to the interviewees for authorisation. The email interviews did not need such editing, which is considered as one of the benefits of email interview (Bampton & Cowton, 2002). Next, these interview materials were read thoroughly in order to highlight substantive elements (Gillham, 2005 Chapter 18), or those elements which are not repetitive, redundant nor part of normal act of verbal communication. Subsequently, when needed, statements were put in a chronological order.

The material was coded and categorised in an iterative and bidirectional manner, both bottom-up and top down. Coding started from the first edited transcript and continued throughout the analysis of all transcripts. The researcher was focused on extracting or deriving categories from the interview material and then applying back those derived categories to the source. In this way, common categories can emerge that form a framework along with additional, different elements. These are then assigned a descriptive and defining name, a process which, according to Gillham, “is a creative business in itself”, and they are going to form key points or “sub-headings” in the findings (2005). This procedure helped to ascertain that new

ideas from the interviewees could emerge that go beyond the interview questions, which were rooted in the literature review and the preliminary discussion with some of the interviewees.

Furthermore, to ascertain the highest quality possible of the analysis the researcher applied four rules evidenced by Yin (1994) as cited in Tellis. They indicate to:

- “Show that the analysis relied on all the relevant evidence
- Include all major rival interpretations in the analysis
- Address the most significant aspect of the case study
- Use the researcher's prior, expert knowledge to further the analysis” (Tellis, 1997)

## **4.6 Constrains**

### **4.6.1 Limitations in methods, data collection and analysis**

Gillham aptly observed some of the implications when choosing research methods: “Good research questions almost point to appropriate methods. At the same time, trying to identify methods may suggest that some of your questions are not particularly well framed.” (Gillham, 2005, p. 5). Similarly, Cooper (2009) noted that the research results depend on the appropriate methods and the purpose of the research. Denscombe (2009) pointed out that these choices can be regarded as constrains, as they influence the overall research defining its scale, scope and planning.

The underlying preferences of this study were to choose adequate and feasible techniques for the purpose of the research. More importantly, the main purpose was to ensure the best environment to the study participants maintaining their relationship and trust, and overcome the physical distance between the researcher and the participants, while leveraging online tools to engage with them.

Gorman et al. highlighted the value of mix methods claiming that they are essential to address research questions expanding the breadth of the research: “competent researchers today realize that confining an investigation to a particular investigative approach does not yield the fullest understanding of a phenomenon.” (2005, p. 12). In addition, they observed that a dual approach can compensate for any drawbacks related to the positivist or interpretivist research paradigm and draw benefits from both (Gorman et al., 2005).

Furthermore, several authors discussed implications regarding the virtual environment, in particular interactive online studies (Bampton & Cowton, 2002; Cooper, 2009; Gibson, 2010; Sade-Beck, 2004). James & Busher (2009) highlighted the importance of the venue in the interaction between the researcher and the interviewee. They claimed that even if the notion of social space is transferred in this ‘other’ reality and, the interaction takes place in two spatial dimensions, virtual space might affect the interview.



Sade-Beck even observed that the virtual world represented a limited picture; therefore, she highlighted the need to immerse in the real world (Sade-Beck, 2004, p. 8).

Reflecting on the interview technique, whilst the main disadvantage of interviews as the source of data is the quality of transcriptions, the other factors being cost and time (eVALUED, 2006), online interviewing shows several negative aspects. Greifeneder even considered them as non pertinent to or usable in traditional qualitative methods (2010, p. 2,7). James & Busher (2009) discussed some relevant issues regarding online interviews, including the commitment of participants and discomfort with video camera. Indeed, some interviewees observed difficulties with setting up special equipment and arranging interviews. Moreover, interview participants may feel uncomfortable with the video camera which can affect their behaviour (2009).

Furthermore, email interview was employed bearing in mind its main disadvantage, namely lack of direct interaction and written, thus more formal approach. This lack of qualifying elements of verbal communication entails from the remote dislocation of interacting individuals, which is typical of virtual communication (Bampton & Cowton, 2002). The research questions were sent at once in the body of the message to give the interviewee time to respond, and this was due to time constraints outlined above. Such approach shares many weaknesses of a structured or standardised interview which is typically used in quantitative research and resembles the survey method (Gibson, 2010). Another risk associated with sending all questions at once was that it could hinder the interactive and informal style of email conversation. On this regard, Gibson (2010) pointed out that this is inherent to any written account.

Asynchronous interaction has another ethical element, as respondents have an opportunity to convey desired information and not their spontaneous reactions under pressure “to protect them from making injudicious comments”. This applies both to interviewees and interviewers, as they are given an opportunity to avoid unwanted episodes.

However, issues also regard online interviewing, such as problems with the Internet connection or a poor quality connection to name a few; they are related to particular software and hardware employed.

The final remark regards the researcher who was involved in a small scale study on research data management carried out at the Loughborough University Library; therefore some elements of observation are also present and this influenced the data collection:

“Participant-observation makes the researcher into an active participant in the events being studied. This often occurs in studies of neighborhoods or groups. The technique provides some unusual opportunities for collecting data but as well could face some major problems. The researcher could well alter the course of events as part of the group, which may not be helpful to the study.” (Tellis, 1997)

#### **4.7 Ethical considerations**

These considerations not only apply to the research conduct, but are relevant when research involves human subjects, which was the case. Indeed, the main source of data were humans and data were mainly collected through interviews; thus the researcher needed to address the ethics of interviewing (Kvale, 1996, pp. 153–157) regarding informed consent, confidentiality and consequences.

Accordingly, the researcher provided all the participants with the information sheet and consent form which were used to inform the participants about the purpose of the research and what was expected from them (see Appendix 2 & 3). All the participants were given the possibility to choose whether to disclose their personal information or keep it anonymous and were informed that the information they provide would be treated with confidentiality and how it would be used. In addition, the data lifetime was limited to the end of this research project, namely June 2012, as this was considered the most appropriate. Furthermore, the participants were given the possibility to review an edited version of the interview transcript and only with their authorisation this material was used in the research.

The ethical procedures were further validated through conforming to the UK Data Protection Act 1998 that constitutes the legal framework in the UK.

Interviews were organised according to a protocol, which consisted of the opening phase, the interview itself and the closing phase (see Appendix 4). In the beginning, the participants were thanked for their participation and were asked permission to record the interview; the recording was only of the voice and not the video recording itself. Then, the purpose of the interview was explained, and the researcher briefly introduced how the interview will be organised. The questions were designed to take no more than twenty minutes to allow the participants to reflect and comment on them. At the end of the interview, the researcher reminded that the participants could comment on any matter if they needed and the researcher expressed her gratitude to all the participants for taking their time and participating in this study.

#### **4.8 Trustworthiness of research – Value of research**

Case study method is often judged critically for not adequate or unclear methods used along the research process (Gillham, 2005). As reported in Tellis (1997) Yin (1989a) pointed out that the applicability of a case study depends upon methodological qualities and rigour “general applicability results from the set of methodological qualities of the case, and the rigor with which the case is constructed”. To avoid such ambiguity and to obtain a rigorous and thorough investigation of the case this investigation applied the criteria of transferability, credibility, dependability and confirmability, which were regarded as appropriate for the chosen, qualitative research method (Pickard, 2007, p. 18).

#### **4.8.1 Credibility**

The criteria, which can be applied in this research, are prolonged engagement with the research participants, triangulation of the techniques used and member checks. Researcher established contacts with potential participants during the internship period in November and December 2011, and this allowed to have first conversations on topics relevant to the study and to identify the potential participants. Also, the study employed two different data collection techniques, interviews and documents review, to balance biases and strengths of these techniques. More importantly, the verification process was done with the key informant who commented on the data analysis and conclusions to validate the findings and to ensure that the data analysis did not contain inaccuracies.

#### **4.8.2 Transferability**

Contextual information needs to be described in details and precisely in order to comply with this criterion. According to Denscombe (2009) the findings from one study can be applied or transferred to other cases if the individual context is provided with sufficient detail and precision. Therefore, this study offered all relevant information about the University and the Library providing background to the case under study. In this way, based on the contextual proximity, the findings of the individual case can be transferable to other cases.

#### **4.8.3 Dependability**

Of concern here are appropriateness and relevancy of the methods and techniques that are part of the research design and an audit by an external person. The choice of methods and techniques is obviously subject to an individual bias; however, their appropriateness depends on the specific aims of the inquiry, as discussed previously. This inquiry sought to develop an understanding of roles, responsibilities, both current and future in research data management particularly aiming to get an insight into approaches and opinions of some of the stakeholders involved in this process. Therefore, interview was chosen as the main technique along with documents that were judged relevant to gather in depth information.

Audit is a means of ensuring that the research is based upon the data obtained in the data collection in order to validate the final outcomes (Pickard, 2007, p. 20). Whilst due to confidentiality of the interview material this criterion was not applicable, this aspect is especially relevant to the topic of this investigation.

#### **4.8.4 Confirmability**

Confirmability aims to validate the results and establish if the final outcomes are drawn on the data that have been collected and analysed for the purpose of the research. Similarly as with the audit, this issue is essential to the topic of this investigation. Whilst the data are not attached to the final work, future studies may change this approach.

Other concerns regard subjectivity of the research. Gillham made a point that interviews are built upon subjectivity which is the foundation of relationships between human subjects “In an interview, the interviewee is ‘constructing’ themselves in what they say, of course, but so also is the interviewer. Inter-subjectivity is at the heart of all social relations, whether in a research context or anywhere else. Acknowledging this does not mean that we get lost, as researchers, in a welter of subjectivity, rather, that we have to consider the role of this dimension.” (Gillham, 2005, p. 6). Thus, inevitably research is subjective, and this is acknowledged.

## CHAPTER 5: Data analysis and discussion

### 5.1 Data Analysis

This study aimed to indentify and investigate the current support services, roles, responsibilities and requirements of stakeholders in the implementation of research data management and to explore how they may evolve in the future. In addition, as described in Chapter Four *Methodology*, the data were elicited from the documents and interviews. Therefore, according to these criteria, it was possible to identify a set of categories responsive to the objectives of the study, namely: 1) Current activities, 2) Stakeholders and roles, 3) Current issues, 4) Future developments and 5) Solutions.

These categories are firstly presented and supported by quotations from interviews and documents, then analysed and discussed thoroughly in an attempt to explain them.

---

#### 5.1.1 Current activities

*This section presents activities and support services currently available at the university that were identified as relevant to research data management.*

The answers revealed several themes and a few common aspects. When asked about current services, almost all of the participants highlighted that research data management is in the early stages referring to the current situation using key terms like “early stages” and “early days”. Two interviewees underlined respectively the requirements of funders with regards to research data management [F] and the activities of research data management Working Group [B, G].

The analysis showed that, at the moment, Loughborough University has no formalised structure, however recently they started considering research data management. Consequently in March 2012 a Working Group has been initiated by the Library Research Support Group and set up by representatives from the Research Office. Furthermore, the group conveys a representative from IT Services, the department of Information Science and the School of Civil and Building Engineering.

The results revealed a variety of services, roles and responsibilities that are provided on campus, ranging from physical infrastructure, large storage systems to guidelines on data collection and storage. In addition, there are other activities as well that two of the interviewees illustrated quite vaguely: “at the moment there is some good practice that we could provide” [C] and “we have staff to work on research data management, there is some advice” [D]. Also, frequently interviewees observed that current activities and roles, like training and duties of academic librarians or IT staff are not focused specifically on data management.

One interviewee, although stated that until now the library hasn't had a role in research data management, indicated the Research Support web page of the Library<sup>1</sup> as a summary of their research activities. This is supported by the strategic plan of the research support team, according to which the library has recently developed web pages dedicated to the management of research data, thus starting from February 2012 it has a section on research data management.

Clearly, the results indicate that when developing new processes many of the activities are undecided and subject to change. In fact, the current situation is evolving, and information about services and support was comprehensive as well as incomplete as there are plenty of services and support that are not yet identified as pertaining to research data management. Furthermore, the results showed that the Working Group has been initiated timely to address the mandate of research councils regarding the management of research data and start considering data holdings of the university.

#### **5.1.1.1 Data collection and storage**

The data revealed that the university offers guidance on data collection and storage for specific types of research data, and this is mentioned in a few institutional documents that are pertinent to the wider research community.

The university has adopted the University Data Protection Policy to comply with the UK Data Protection Act 1998 which, in a few words, is the UK legal framework applicable to personal data and information held in automatic filing systems. The policy includes a section on academic research for researchers dealing with human subjects; the main aspects are further outlined in the guidelines provided by the Ethical Advisory Committee. They have developed and implemented the "Code of Practice on investigations involving human participants" and issued specific information about research activities such as data collection and storage covered in the "Guidance notes for investigators. Data collection and storage". This document informs about the secure storage of research data and their disposal, and data are categorised according to the type of data, such as numerical/statistical data and blood samples on which specific information is provided concerning the length of storage, disposal, and if applicable long time storage in case of longitudinal studies or policy-making works. In addition, there are some indications on the responsibilities for research data, in terms of getting rid of the data appropriately and informing about the will to retain the data:

"The principal investigator (or supervisor where student projects are concerned) is responsible for ensuring that data is destroyed and disposed of in an appropriate manner."

"Investigators who leave the institution at which data was generated should obtain permission from their Head of Department to retain data/copies of data. Permission should only be granted where it is clear that future use will be

---

<sup>1</sup> <http://www.lboro.ac.uk/services/library/research/>

consistent with the terms of consent.”

However, there are no further indications on data storage, including where exactly the data should be kept (institutionally provided storage or individual solutions), how the data should be disposed of (secure disposal), nor how the requirements would be enforced.

It is worth noting that so far the abovementioned Code of Practice and Guidelines are one of the few official documents that explicitly deal with research data management. More information on other available services and official documents is presented and discussed in the *Stakeholders* section.

Unofficially the Library has been considering research data at least since August 2011 as a search done on the university web site in February 2012 showed a working draft of the Institutional Repository Frequently Asked Questions, which declares:

“Research funders often require data supporting research to be made open access. We are happy to consider data for inclusion in the Institutional Repository where requested, on the understanding that as this part of the service develops there is a possibility that the data may be moved to a separate repository or store.”

The final consideration is from the technical point of view and the interviewees with IT or computing background and interested in technical aspects gave more details on this matter. They discussed the physical infrastructure, such as servers, the Institutional Repository and identified different types of storage that are currently available: centrally provided, local storage used by researchers and large systems for storage.

#### **5.1.1.2 Ethical considerations**

One interviewee illustrated ethical considerations regarding data collection and storage and, referring to the Code of practice, stated that ethics is primarily concerned with secure storage and disposal of data and observed that: “Ethics [looks at data] from Data Protection act side, and from where it is going to be stored, so we check that data is stored securely, but we don’t really go into how long they want to keep it or why they want to keep it.” Furthermore, the ethical point of view was summarised as: “essentially at the moment the ethics policy is after a certain number of years destroy data” and the ethics point of view is the consent that is agreed upon between the research participants and the researcher that states “we will be the only people who see the data” [B].

In addition, one interviewee mentioned the only example of data deposit that has gone through the ethical committee, and this is supported by the Minutes of the Ethical Advisory Sub-Committee, which is part of the Research Office. One researcher wanted to deposit data in the UK Data Archive and the Committee required information related to the specific archive in which the data would be deposited, how the data would be made anonymous and that the study participants would be properly informed about data reuse.

“Now ‘put data in the archive and lots of people can ask permission to use it’; it’s making sure that people fully understood what would happen to the data, would be hard to fully anonymise data (personal memories). Researchers need to be made aware of it in terms of ethics. It’s a change of direction in what you are saying to the people, especially in the social sciences who ask personal questions and that’s the data that seems to be wanting to put in an archive.” [B]

This key focus on storage and disposal of data is confirmed in the documents that researchers are required to comply with, such as the project completion and research proposal form for human, biological or psychological and sociological investigations, which require information about the storage and disposal of data. Similarly, the Guidance notes for researchers lay the ground for research work in terms of secure storage, backups and copies of research data and described it as follows: “data should be stored safely with appropriate back up and contingency plans in the event of loss, damage or unauthorised access to the data. Wherever possible a complete duplicate set of the original data should be retained.”

These results indicate that specific guidance and services are provided to the research community even without a central structure. Furthermore, the university and especially the Working Group has already started considering a more coherent approach to research data management.

#### **5.1.1.3      *Research computing***

From the responses, it is also clear that there are specific data services for researchers doing computational work and this was highlighted by some interviewees who mentioned Research Computing services. In particular, this service ranges from the campus High Performance Computing to local research workstations and large storage system:

“Our current main tasks are to support the campus High Performance Computing (HPC) service and [...] we also have a wider role of involvement with any research computing activities. As a primary point of contact for those doing computational work we maintain some large research data storage systems (150 TB) and have some sight of what researchers are doing, e.g. on local workstations, or where researchers put data when they remove it from the HPC systems.” [E]

#### **5.1.1.4      *Researchers’ data practice***

A few interviewees provided information concerning researchers ranging from their behaviour to requirements. Two interviewees gave divergent opinions on whether the Institutional Repository or the Research Publications Database contained research data:

“The slight issue with institutional repository is that people tried to put data into repository, because it's a place for them to store it. The word ‘repository’ makes it sound it’s a silo for everything. They [the staff] had issues in the past of people trying to put all of their study data.” [B]

“The only data was that associated with an article, i.e. graphs probably because they have submitted a document without charts. Each department was given a half-hour training session or had the opportunity and it was stressed that the



purpose of LUPIN was to record research outputs, publications as opposed to research data, except when data itself was being published which occasionally happens.” [C]

Furthermore, one interviewee stated that some researchers expressed their interest in publishing survey and questionnaire data:

“I was asked about questionnaire and survey results from a researcher and it’s very rare that raw data from questionnaires and survey to be included in the Institutional Repository or publications databases, but that’s the sort of thing that people are starting to request. Questionnaire design is quite a hot topic and there is a lot of concern about it, because constructing it is an art and some people are wary of making it available because there is lots of intellectual input.” [D]

These results together with the previously mentioned example of research data deposited in a data archive indicate that there is some interest in depositing and publishing research data, which is already supported by the university and in particular the Research Office, Ethical Advisory Committee, Academic Librarians and the Institutional Repository staff.

A similar interest in publishing and providing access to research data was indicated in the report of a recent investigation on researchers’ views on Open Access that was undertaken at Loughborough University and in particular two comments pointed out a few aspects regarding research data:

“I read on the EPSRC website that in a few years time all data from research projects should be publically available - what's the university view on this?”

“Creation of open access database not only for articles or chapters from the books may be important data, values, processing protocols at the university level will be really helpful.”

---

### **5.1.2 Stakeholders and roles**

*This section covers specific roles and responsibilities of support services that were described as involved in the implementation of research data management, conveners in the RDM Working Group or potential partners in the implementation and provision of research data management.*

The interviewees identified various stakeholders and their roles in research data management. The most common were researchers, academics, the Library, Research Office and IT Services who are already convened in the research data management Working Group.

A few responses mentioned the university managers, senior managers, the schools and more specifically associate deans for research, thus emphasised the needs and interests of the university. However, some interviewees listed also the Enterprise Office, the library systems team, Graduate Office and the facilities management. In addition, one interviewee mentioned research records keepers and archive services, who, though, can be recognised as the Library and the Archive but this was not clarified.

Regarding the research data management Working Group, it is worth saying that in March 2012, when the majority of the interviews were held, the group had no formalised structure. This was noticed by one of the interviewees, who stated the following: “we are trying to persuade some academics to sign up to this group as well”. [G] In the end, according to the information from a Loughborough University webpage dedicated to Open Research Data (Loughborough University, 2012c), the Group consist of the following people and support services:

“Angela Crawford, Ray Kent and Zoe Stockdale from the Research Office, Jeff Brown and Lizzie Gadd from the University Library, Martin Hamilton and Julian Highfield from IT Services, Marcus Enoch from Civil & Building Engineering and Adrienne Muir from Information Science.”

It is worth noting that the management of research data is a joint effort and brings together different services, roles and responsibilities. These are mostly sustained by central support services that are relevant to the wider institutional community and support strategic aims of the university, as well as other entities that have a stake in this and also need to be taken into consideration. The contribution of each of the department was illustrated as follows.

---

#### **5.1.2.1      *Research Office***

Part of the interviewees attributed a special role to the Research Office. The data revealed that they were recognised as the institution most closely connected with researchers and their role was described as a first and a central point of contact and a one stop place. In addition, some interviewees observed that staff of Research Office was aware and equipped to support RDM.

The contribution of the Research Office was thoroughly described by one interviewee who illustrated their current involvement in various stages of research activity and envisioned their future role in research data management:

“effective introduction of RDM support will require involvement at the points of applying for grants (due to cost implications), grant mileposts (for recording deliverables) and publications records (to attach data records to related papers) so will require support from the Research Office and from associated business processes.” [E]

---

#### **5.1.2.2      *Research Support of the Library***

Recognised a similar role as the Research Office in supporting the research activity, some interviewees described the Library as best qualified to provide RDM, even if divergent opinions emerged about whether the library was ready to support it.

The contribution of the Library is outlined in their strategic plan for 2010/13 which indicates that they will provide optimum support for the research. In the first instance, it is clear that the Library is mainly focused on information sources, which are mentioned as one of the library values. However, the first goal

is clearly research-focused: “to develop and deliver optimum support for research and enterprise activities of the University”, and will be achieved through the knowledge of researchers’ needs, by aiding researchers in the discovery of resources and assisting the university in promoting research outputs. More specifically, this is done through liaison activities, providing infrastructure (institutional repository and research and publications information system), support and services for Research Excellence Framework (REF).

Furthermore, the only document that mentions research data management in an explicit manner is the operational plan of the research support team for 2011/2012 that reflects the strategic directions of the Library. The research support team is responsible for activities which are outlined in their strategic plan and are built upon understanding research practice and researchers’ needs. They include establishing professional contacts, regular surveys, working closely with researchers, and delivering tailored teaching in information literacy. More specifically, they support the entire research lifecycle, including funding and experimentations, keep abreast of changes in the research environment and adapt to them. They are focused on informing colleagues on best practice. Also, they develop and provide tailored teaching that is primarily focused on information and management capabilities. The role of this team is also supported by other documents, namely the individual library wide remits and cross-library teams that outline their roles and responsibilities in a similar manner.

---

### **5.1.2.3 Academic librarians**

The role of academic librarians was discussed and underlined by a few. They were described as key people with regards to research data management given their relationship and engagement with researchers, in particular through such activities as teaching and participating in research meetings:

“We’ve seen librarians as key regarding the teaching processes, that’s strength of Loughborough University and librarians work hard to build these contacts.” [C]

“We try to get on student staff meetings [...]. I’ve been to quite a few Information Science presentations and some Computer Science and they are quite surprised to see us there.” [D]

The findings are supported by the documents as, according to the job description of academic librarians<sup>2</sup>, their role is “to deliver appropriate library services”. At present, they are mainly focused on information needs and only eventually are responsible for other duties that may be required from them. This implies that their duties might change in line with the needs of users. In addition, they work together with the library liaison officers to offer the right level of support to researchers, staff and students.

---

<sup>2</sup> The researcher was informed by one of the Loughborough University Library staff members in the period between November and December 2011 during her internship that the job description would change in the nearest future

Finally, one interviewee mentioned the role of academic librarians during the Loughborough University research staff conference in 2012 to promote their role and activities by encouraging researchers to get in contact with them and mentioned the topic of research data management: “People were also encouraged to contact the academic librarian, 23 did and part of the follow-up emails included research data management.” [D] This is supported by the report of the librarian that described promotion and advocacy during this conference in its previous edition.

---

#### **5.1.2.4 Primary Responsibility**

Discussing who would hold the prime position and was best qualified to support research data management interviewees provided divergent opinions. Some participants stressed that researchers would go to the Research Office or Graduate School rather than to the Library by describing the Research Office as the main reference for research:

“I don’t know if the library will be the central location, we’ve got the web site. Most people would go to the Research Office or Research Student Office [...]. They will be the main contact when it comes to any research.” [D]

“The library is best qualified, just many researchers won’t use the library as a first place. Research Office and Research Student Office are used as central point of contact.” [C]

However, as stated above, they also recognised the importance of the Library, and this is clearly reflected in a consideration of one interviewee who stated:

“In all likelihood, the Library – not the Research Office - will have the primary responsibility for oversight of RDM.” [H]

Mostly, the interviewees observed that provision and oversight of RDM would be a joint effort and, as stated previously, three key stakeholders were collectively identified by the interviewees, namely the Library, IT Services, and the Research Office. They were described as “interested units” that would collectively provide and coordinate the RDM processes. This was briefly outlined by one of the interviewees:

“A combination of the people: Research Office, Library, IT because they come from different angles, can give different but complementary advice.” [B]

Furthermore, some interviewees distinguished between front and back office roles, or who would engage directly and indirectly with researchers. One interviewee stated the following:

“A one stop place in terms of why we are doing it will be the Research Office because it's going to add up money to your pocket. The library can contribute of how the best way to archive, the actual process and what you need to archive properly and make sure that data can be retrieved, [...] IT is the physical space for the data, an indirect kind of link.” [B]

A certain difficulty to indicate who would hold the central role was reflected in other considerations. For

instance after the first meeting of the group, one interviewee, when asked about who would be involved in the Working Group, mentioned that they are only “starting to think about it” [G].

Finally, one interviewee emphasised the importance of collaboration between support services and highlighted their contribution to the strategic value of the university:

“I also see significant benefits in having the different support services work together on a topic of strategic value to the University.” [H]

---

#### **5.1.2.5 Collaboration**

The results also revealed that the Library has a particular focus on collaboration within and beyond the university. This is articulated in several documents, ranging from the report of the librarian to the strategic plan of the research support group, the individual library wide remits and cross-library teams. For instance, the report of the librarian for the period of 2010/2011 highlighted a close relationship between the Library, academic departments and support services, namely the Research Office, the Graduate School and a more general network of professional contacts “to ensure that Library services contribute to the learning and research of the University”. The Library aims to be a main liaison point for support service departments and also achieve their goals that can be reassumed in the following terms that emerged from the library strategic plan and report of the librarian, namely: collaboration, communication, training, and promotion of the library support services.

Furthermore, a few interviewees mentioned working together with external partners, including the Digital Curation Centre (DCC), local higher education institutions (HEIs) and in particular the local group of research led universities.

According to a few interviewees, the involvement of the DCC was described differently as an institutional partnership to develop a coherent approach to research data management, a partner that would provide training sessions and prepare academic librarians to offer research data management support.

Another significant finding is about the standard approach to research data management. Although one interviewee pointed out lack of standards for implementing RDM, most of the interviewees indicated, when asked about the services, training or tools that are required to provide research support services, materials and tools developed and provided by the DCC.

Also, some interviewees acknowledged that a significant part of the institutional research is done in collaboration with commercial partners or highlighted their importance and role. One interviewee observed that:

“Where research data are not wholly owned by the University, the other joint owner(s) will be consulted over how best to manage jointly-owned data.” [H]

Some interviewees hypothesised a further collaboration with local HEIs and more specifically, two interviewees referred to the collaboration between the Library and other local academic libraries and another mentioned a local research support group.

This is also supported by the report of the librarian that describes the East Midlands Research Support Group in which the Library had a leading role and which aimed to share best practice in research support and other activities by means of online tools. However, no further information was available.

“If we found that the major capital investment was the facility, it would make sense to do something collaboratively. We’ve been already talking about the equipment sharing mandate. It follows that there may be other conversations about collaboration; maybe research data is one of them.” [G]

The last comment is about a recent collaborative project aiming to enhance the provision of research computing and in particular large storage systems. According to the High Performance Computing Midlands webpage (HPC Midlands, 2012) Loughborough University and the University of Leicester with support and funding from the EPSRC initiated a new service that provides cloud supercomputing for academic institutions and industry.

---

#### **5.1.2.6 Research Councils**

*This section presents the requirement of research and funding councils and other agencies that were identified as one of the major stakeholders when implementing research data management.*

A particular attention was dedicated to research councils and other external agencies that recently set a requirement to manage research data that are funded by these agencies. More details on this topic can be found in Chapter Three.

Most of the interviewees collectively stressed the importance of research councils to the university, in particular, the expectations of the Engineering and Physical Sciences Research Council (EPSRC) to develop a roadmap on sharing research data and subsequently implement it, respectively by May 2012 and be implementing by 2015. Furthermore, two interviewees pointed out their role, as EPSRC accounts for 50% of the university research income and Loughborough University is a Framework University for EPSRC, underlying that firstly they would focus on EPSRC mandate and subsequently would need to adapt it to the general research practice. This is supported by the information that can be found on the EPSRC web pages that outline the agreement between the funder and the institution (EPSRC, 2012).

Furthermore, two interviewees described this mandate as a stimulus to improve the current practice. An additional comment, pointed out by one interviewee, was that research data management plans are intended as best intents and would evolve and potentially be re-use:

“Asking what our research plans are may provoke further work: some projects funded to provide case studies to help

prepare institutions or even suggest a generic plan that other people adopt; maybe refining, sharing or publishing some plans and revising them later on. The plans are ‘best intents’ and can be changed. It’s getting people think about research data” [C].

From another point of view, one interviewee pointed out drawbacks and a contradiction between the requirements and what is done in practice:

“I have got an impression that Research Councils have given an enormous number of guidelines to universities. Research has to be done as a result or condition and you are required to data manage the project, it’s often like your university will let you know what you need to do, sometimes they may not.” [D]

Likewise, one interviewee stressed the different and contrasting requirements of funding councils and other agencies:

“NHS could be telling the researcher one thing, and the funding council another and [the Research Office] might need to start negotiating” [B]

---

### 5.1.3 Current issues

*This section presents the views on the current situation that emerged from the data and were described or related to in such terms as “difficulties, issue, different, problem”.*

The results revealed a common understanding of some major issues, like lack of central coordination, research data practice, cost, researchers’ approach to data. These were summaries by one interviewee in a very straightforward way:

“There is no central contact point for RDM issues; hence no institutional oversight of RDM; nor has an institutional policy for RDM yet been formulated. The University is not aware of its research data holdings, nor where its researchers have placed their data sets (e.g. in a national database or service). The potential for re-use of data has not been considered. No provision exists for large-scale (petabyte) storage of research data.” [H]

Almost all of the interviewees recognised that there is no central approach or coordination of research data management, which was illustrated as: “no coherent approach”, “nothing centralised”, “no central coordination”, “no institutional oversight”, “little consensus” and “no formalised structure within the university”, and “lack of central support”. In addition, three interviewees recognised the need to have an institutional policy regarding the management of research data.

Two interviewees explained this situation as lack of the requirement or recognition from the central university; others considered this as an unrealised potential:

“A lack of recognition of the business need for RDM.” [E]

“As far as the central university is concerned, they have not been particularly interested. Now with the planning, they know, it raised the profile of research data significantly.” [C]

Finally, it is worth noting that the considerations about the lack of institutional oversight and the requirement of research councils are in line with what is currently happening at other HEIs in the UK and beyond that are starting to consider RDM. Whilst some of the UK institutions obtained financial support to handle RDM, Loughborough University has not obtained such funding, yet they showed capacity to approach and develop these new processes.

#### **5.1.3.1 Researchers' data practice**

Another aspect was jointly discussed by most of the interviewees: research data practice and, related to it, researchers' approach to research data and support services.

The data analysis showed that research data management is done by individual researchers, whose practice varies dramatically. This was evidenced by using the following terms: "on their own basis", "on an ad hoc basis", "insular" and the situation was described as "business risks, haphazard data storage practices" [E]. The data storage and management was also illustrated as "that golden PC which has all of the research projects data on it" [G].

The majority of the interviewees expressed their concern about the approach of researchers to the research data, who are apparently not aware of the value, quality or quantity of the data they produce, nor recognise the need to know about it. In addition, researchers were recognised as the main responsible for the data:

"Persuading researchers that this is important and that there are benefits to them."

"Recognition from researchers that this is an essential part of their workflow." [A]

"The lack of assessment by researchers of the quantity, quality and value of the data they produce" [E].

"[They] are not asking the question because they don't really know that they need to" [B].

When discussing the research practice, a few interviewees expressed concerns about data loss or invisible data. They shared a common view as they claimed the risk that researchers may leave and take their data with them:

"When the individual leaves, this can be a problem in terms of research data that they collected. Research data is either on the central systems or hard drive, [etc.]. We've probably lost enormous amount of research data, because somebody has taken it with them." [G]

This issue is supported and in part addressed in the "Guidance Notes for Investigators Data Collection and Storage" which states that researchers who leave the institution hold the responsibility to inform if they want to retain data:

"Investigators who leave the institution at which data was generated should obtain permission from their Head of Department to retain data/copies of data. Permission should only be granted where it is clear that future use will be consistent with the terms of consent."



One interviewee stated that this issue is and can be potentially addressed by the Library that could prevent data loss:

“It’s important to meet them at the beginning, but there is a need for some ongoing support and perhaps checking towards the end, when that group may disperse; at the moment we’ve got no requirement to do so, sometimes it happens but it’s not planned.” [C]

Frequently, the responses showed that the ownership of research data is an issue. On one hand, interviewees recognised that research data are now of interest to the university and to the wider community. However, two interviewees recognised that there are issues with shared ownership when research is done with commercial or third party partners. [quote]

A few interviewees pointed out lack of recognition of the Library’s role and mentioned some reasons that influence the approach of researchers. According to one interviewee:

“We often find that established researchers do not think of the Library as a department that can help with their research. So there is an issue in promoting our services to this group.” [A]

The approach of researchers was illustrated in a straightforward way by one interviewee:

“At the moment researchers are puzzled and bemused by our interest in their research data – it is almost as if we’re offering to help them organise their desk drawers!” [F].

The issue of cost was mentioned several times and with different facets. The results revealed that there is a significant cost of storage for research data, including the general cost of infrastructure [G] the storage of research data [F], “the potentially high costs of long-term data storage” [E] and the cost of metadata and long term curation [A]. One interviewee illustrated the issue of storage and mentioned that currently researchers need to sustain the cost of storage on central services that may not be adequate for research that generates volumes of data: “Research groups can ask for a shared drive on central services, though there can be significant costs for them as the volume of data grows.” [E] In addition regarding HPC, there are significant costs of providing back-up storage and partly for this reason the offer of storage for HPC is limited: “high performance computing storage systems are purely for transient (working) data and users are expected to store copies locally” [E]. Besides, one interviewee lamented that there are no real economies of scale [F]; however this differs from what other responses revealed. More details are provided in the *Solutions* section.

#### **5.1.3.2 Support services**

Furthermore, the responses showed a shared awareness among the interviewees that support services are not providing adequate services to researchers. Some people highlighted not only the lack of infrastructure, such as hardware and software, but also the lack of skilled staff or that staff is not yet equipped. For instance, two participants described respectively that the Library staff was described as

currently not equipped to support research data management or not specifically focused on this, and the IT staff is not tasked to support this process. In contrast, the staff from the Research Office was considered as equipped to support research data management if provided with training.

One variant concerning researchers emerged from the data, when support services were depicted from the point of view of researchers and described in the following manner: “slow moving bureaucrats would never do anything to address some of data problems” [G].

Finally, one interviewee dedicated some attention to defining research data management, which was opposed to ethics and described as an issue: “Ethics is sort of against research data management policy: use it for one study and get rid of it whereas data management is: use it for lots of studies and share it” [B]

---

#### **5.1.4 Future developments**

*This section presents events and activities that were described collectively as evolving and work in progress and were identified in the strategic documents.*

The responses showed several activities and the following are the major developments as depicted by the interviewees, as well as other significant findings.

One interviewee presented and reassumed in a concise manner future events and activities that are framed by issues such as central control, policy and inadequate storage for large data that were illustrated in the previous section:

“We are committed to investigating all of the above issues (and more!) over a time span of up to 3 years, during which we envisage that certain RDM services will be set up in order to fulfil the needs of the University and the requirements of external agencies. The nature and scope of these services will ultimately be determined by senior management, drawing upon advice from the RDM Working Group and others.” [H]

In an additional comment, related to the possible future evolution of the Working Group, one of the interviewees envisioned a steering group that would provide advice and central management and would involve departmental administrators along with other stakeholders to work together:

“Once we have get going the service, we might establish a steering group, which will provide some advice and central management, and consisting of representatives from the Library, IT services, Research Office and academics, possibly departmental administrators and members of the research committee” [C]

---

##### **5.1.4.1. Working Group**

Activities of the research data management Working Group were most frequently mentioned. In particular, a few interviewees agreed on the main topics that the group will address, such as the institutional policy and education.

As showed in the results, the institutional policy on research data management will need to comply with the requirement of research councils concerning research data; particularly relevant is the mandate of EPSRC to have a policy on research data and then an implementation of this policy, respectively due in 2012 and 2015. Its aims were clearly stated by one of the interviewees:

“A policy that helps researchers know what they have to do with their data, what the research councils are looking for, tailoring it to our research”. [B]

Regarding training, the responses indicated major directions how education and courses may evolve and be more research-focused. In particular, this would involve advice and training provided by the Library and the Research Office. One interviewee envisioned that: “All new researchers should attend a training course on RDM.” [F] Likewise, another person expressed that there is a need to have some guidance on research related topics: “handbook of managing research projects” [G]. However, it is not clear what would be the content of such handbook and whose responsibility would it be to deliver it.

---

#### **5.1.4.2. Academic Librarians**

A few specific considerations were provided in the responses on the role of academic librarians. One interviewee highlighted that there is a need to increase research related activities if academic librarians are supposed to provide support to researchers and observed that this would require changes in the current duties:

“If subject librarians need to be that main contact points between departments and the library for research, we need to scale up the research, increase liaison, but that’s time consuming. Not only promoting research but also looking at where we can shape time for other activities, and the consequence for subject librarians is to drop something else” [C].

Some interviewees identified academic librarians as responsible for filtering questions coming from researchers and building a database of answers and expertise. Their role could also include the promotion of research data management through the research blog and sessions with researchers to present their support.

---

#### **5.1.4.3. Institutional Repository**

A few interviewees considered future developments of the Institutional Repository. It was described as storage for any type of content ““repository makes it sound it’s a silo for everything” [B] and another interviewee clarified its role in providing access to research data: “If I submit a publication to a journal [...], you should have the supporting data and you should be prepared to make that available; the repository model is a good way of doing it” [G].

One interviewee envisioned “a repository of metadata about research data on campus and the development of “a ‘joined-up’ database of research equipment (wind tunnels, anechoic chambers, etc), the resulting research data, and the research outputs (journal articles etc)”. [F]

Besides, two interviewees described the institutional repository as a standard and underlined that the university has recently introduced a mandate to deposit research outputs, namely articles, conference papers and book chapters, and this is also outlined in the report of the librarian:

“Other stuff which is much new, like the IR; we’ve only recently introduced the idea that everyone should deposit in it, that was quite a big change for us, the mandate is fairly new.” “Institutional Repositories, which are relatively recent, are set up and managed in a fairly standard way” [A].

---

#### **5.1.4.4. IT Services**

Moreover, additional information was provided regarding future development of the research technological infrastructure and with a special attention to the role of IT Services that will be increasingly focused part on research data management and they will aim to give researchers “a way of having data stored, replicated, mirrored and backed up” [G].

Two interviewees highlighted the activities and role of the IT Services, which would aim to build a research collaboration service “to help support multi-site, multi-organisation research activity” and there is a possibility to have a versioning archive, which could also support “managing project deliverable and other data”. [E] Furthermore, IT Services would provide the infrastructure and daily support for researchers and they may also help to set up guidelines, and advice other support services, in particular the library “IT might want to come to advice us on what we should say but probably they won’t have any more involvement” [C]. In addition, two tools currently used were identified as adequate for research data, a publications management system and the system used for the institutional repository.

Furthermore, the Library Systems Team will also run the service, install and manage. They will experiment with the Institutional Repository and will put some data sets. In addition, they might consider multiple repositories for research data. On the contrary, another interviewee observed that they would rather not provide repositories according to the type of content and stated that there would be an experimental repository. Only one interviewee mentioned that IT Services would provide advice:

“- Advice on available research data archives and alternative means of data storage on campus – from IT services.” [F]

---

#### **5.1.4.5. Issues**

Some interviewees identified issues with regards to future developments and the results showed that the most common were the education of researchers and raising awareness and the involvement of

researchers in the provision of research data management. One interviewee observed this difficulty as follows:

“There is also the question of how you get the people to come along, whether there are some mandatory training sessions, you can take advantage of ‘you will learn about research data management, people will complain but they know that they have to be there.’ [G]

One common aspect was raised by a few interviewees concerning the need to decide on the content for institutional storage and preservation, namely the research data that are already available, their size, and eventually the data that is available for re-use and can be shared. With a certain emphasis, this was illustrated using terms such as “challenge”, “conundrum” for those involved in setting up research data management. This was clearly reassured by one of the interviewees who stated:

“There is a particular question about how do you find out what is there right now?” and “How do we persuade research active academics and research staff to tell us what they have?” [G]

Whether research data management is a new process or an additional layer, the data highlighted that most of the interviewees had slightly different opinions on this. One interviewee envisioned the possibility to recruit someone for a new role based in the library and a few interviewees noticed that research data management is a change of direction. This was directly expressed by one interviewee who stated:

“It is very different from an ethics point of view, which is the consent [...]. It’s a change of direction in what you are saying to the people, especially in the social sciences who ask personal questions and that’s the data that seems to be wanting to put in an archive. Essentially at the moment the ethics policy is ‘after a certain number of years destroy data’ and that’s written in the code, in the guidelines; will have to say, unless you want to put it in same kind of archive; that is essentially the issue with data management.” [B]

Moreover, one participant envisioned that ethics would need to be tailored to research data management; in particular a number of institutional documents and templates will have to be adapted to the new developments:

“Ethics will have to be rewritten in light of change for data management. Will have to change the templates, everything will have to be rewritten to some degree.” [B]

On the other hand, a few stated that what is currently available would be adopted to take on additional research data management elements. This was supported by such examples as current staff, who with additional training and increased focus on research, would be equipped to support the management of research data and, on the technical side, currently used tools such as the Institutional Repository and the Research Publications System might be implemented for a data repository.

The final comment is about an example of particular data sets, such as outputs of supercomputing and other types of research data. One interviewee pointed out that complicated datasets may exclude the possibility to process them automatically. This was stressed in the following manner:

“How do you know what constitutes the research data output? At some point there is a judgement” [G]

Moreover, the interviewee identified some grey areas and unexplored issues, such as providing access to and making available the research data that were funded by tax payers.

---

### **5.1.5 Solutions**

*This section present one particular aspect of future developments that was identified in the data and that regards solutions for how best engage with researchers and means of how (best) implement research data management.*

The responses revealed differentiated aspects, and it was not possible to identify a shared understanding, nor the questions looked for it. Rather interviewees provided solutions from their perspective, considering professional and cultural background.

Collectively, the interviewees agreed that support services need to give solutions to the current issues and one interviewee identified this aptly: “it will be a matter of not just giving them 25 more things to think about, but 25 solutions to potential problems” [F].

---

#### **5.1.5.1. Management**

Two interviewees highlighted the need to have a shared knowledge and understanding, and in particular they stressed the role of leaders and that they had to agree on a common message:

“drawing a line between what’s meant by data management, data protection, outcomes and outputs project, make sure that everybody knows what we are trying to say [...] communication, being at the same point and share good practice” [B]

“Good leadership at all levels; people who know what they wish to achieve in relation to RDM – and who are given the necessary institutional support (including investment in human and financial resources) with which to achieve it.” [H]

---

#### **5.1.5.2. Incremental approach**

Another theme was common to the majority of the interviewees, namely an incremental approach to implementing research data management and building on what is already in place. This was clearly expressed by one of the interviewees who stated:

“I think the key element is to work within existing work and data-flows as much as possible. Not to introduce lots more processes, but to tweak and enhance existing ones and to make it all as intuitive and seamless as possible.” [F]

Similarly, others acknowledged that there is room for additional information on research data management in the current training programme. One interviewee envisioned research data management as another strand of what the academic librarians should be doing. Also from the library perspective, one participant stated that collecting, indexing, making available and preserving research data is only one step further. This implies that research data management is mostly seen as an additional aspect of the current responsibilities.

---

#### **5.1.5.3. Academic librarians**

As already mentioned some interviewees recognised the role of academic librarians and identified a few specific means to reaching users which were described as follows:

“I still think that the best method is word of mouth. We might reach just one person from a research group, who will spread the news much more efficiently than the in a formal lab group. If you get the results with one researcher, the same with students, that message will get around.” [D]

“Through academic librarians acting as front of a house, that’s the best way forward. They need to be an account manager: collecting information, questions and checking up the answers among IT, RO, other colleagues, and possibly the unions.” [C]

“We have academic librarians who are attached to each of our schools, would be quite natural that they would sit with the associate deans for research to see what they got.” [G]

Some of the identified solutions are already in place at the university, namely to meet student staff and researchers during research meetings. In addition, such involvement was further illustrated and academic librarians could participate in the departmental ‘get together’ to ask questions, as pointed out by one interviewee, like: “What are you doing about storage of your data?” [C]. Furthermore, the means of reaching researchers were to provide information at the point of need and distribute it using available resources, like the library research web site or, possibly, through an advice sheet.

Another relevant finding is the need to meet researchers at the beginning of their career at the university or approach research active academics that had a publication in the last years to inquire about their data. On this regard, one person highlighted that the library might have a role at the end of the researchers’ time at the university. This attempt was also illustrated as working with defined communities such as supercomputing, or researchers who consider themselves as a community. One participant has foreseen getting research groups together and setting up a mailing list.

---

#### **5.1.5.4. Researcher engagement**

Furthermore, some participants focused on the way in which researchers might get involved in the implementation of research data management and researchers, academics were listed among the stakeholders directly involved in the working group or in the future steering group.

One interviewee was the most comprehensive on this regard:

“Using the researchers to see how it will work, how feasible, and probably ask more researchers through a trial, survey to see if it is going to work for everybody. Working in this way seems quite good, you get quite a good feedback. People tend to respond because we asked these people before we rolled out. Get people on board early who will have to do the work.” [B]

On the other hand, interviewees pointed out that some methods were to avoid, namely the following terms were used: “impose”, “inflict” and “dictate”, and one interviewee described this approach comprehensively as:

“The worst thing is to set up a rule without anybody talking about it before” [B].

Finally, it was highlighted by two of the interviewees that the university would look at cheaper solutions for storage that could be achieved through collaboration with local higher education institutions. More specifically, the example described one or two institutions that could provide hardware with a mirror for back-up and preservation. This would enable local economies of scale regarding hardware as well as joint support by means of workshops and meetings. The cooperation between local academic libraries was given as an example, and this was said to be fairly common.

---

#### **5.1.6 Value of managing research data**

*This section presents the findings that described and discussed the reasons why research data management is important and why does it matter to manage research data.*

Responses showed various significant aspects and a common approach, although not all of the interviewees covered this theme. Most frequently, the results showed a dual approach from the point of view of researchers and the institution, such as the value of research data for the university and its enterprise and, in an opposite and parallel way, conceptualised the value of research data as intrinsically part of research practice:

“I know that it is possible view the laboratory notebooks of researchers from a century ago and check and reuse the data from their experiments. We need to ensure that it is possible for people in 100 years’ time to check and reuse the data from research that is happening now” [A]

“We were routinely sharing research data all the time. [...] it was completely accepted to validate the assertions you were making, [...] share the data sets. I think there is a psychological hurdle to overcome, if you are not used to sharing”. [G]



“Research data is both working material and final product of one of the University's major business activities.” [E]  
“Would be helpful, especially on a reporting side, to have a centralised knowledge in terms of data management and data collection” [B]

It is worth noting that two of the interviewees have respectively a background in science and an experience in research, thus their approach might be related to the professional and cultural background.

## **5.2 Discussion**

In this section, the categories that emerged from the data analysis are presented and discussed in the following order: 1) Current activities, 2) Stakeholders and roles, 3) Current issues, 4) Future developments and 5) Solutions.

### **5.2.1 Current activities**

The results revealed that currently Loughborough University has no formalised structure regarding research data management and this situation was commonly described as early stages. Furthermore, the results showed that there is a shared and reciprocal awareness among the interviewees about roles and responsibilities. Also, additional information in this matter emerged.

It is worth noting that not all the participants identified support services that are currently provided and can be defined as relevant to or examples of research data management. Also, as some acknowledged, there are a number of activities happening on campus that were not known to them. Therefore, not all of the available services could be quantified.

Whilst the university has not approached this topic formally and even if they only set the ground for further work, the interviewees showed remarkable attention to research data management, interest in the needs of researchers, and thoroughly identified and discussed aspects relevant to the process; thus the data revealed a rich picture and a number of activities on research data management.

In addition, the interviewees expressed heterogeneous views on current services due to their affiliation and background, and their considerations were not limited to their own services and activities but also discussed those of other departments, namely central support services.

As showed above, current situation implies that research data management is a comprehensive concept including a variety of support and services which at the moment apparently have no shared aspects or definition. The only direct attempt to define research data management is discussed in the *Current issues* section.

Furthermore, from the analysis seems that a few elements were not covered, such as information about the various institutional standards, guidelines or strategies, mentioned only by one interviewee. The reason

that interviewees did not discuss the existing institutional standards may be the distance between the practice represented in the considerations of interviewees and the planning and strategy which can be found in the documents. Also, they may not have been identified as relevant to the topic. Besides, none of the interview questions addresses specifically institutional documents.

The last conclusion is about training, advice and other activities on campus. The responses showed that the majority of activities and roles, such as training provided by the Research Office, Graduate School and the Library and duties of academic librarians and IT staff, were not specifically on research data management or not focused on it. No information was provided whether IT Services or Research Computing deliver courses and such information was not covered in the documents, neither it was discussed with or mentioned by the interviewees.

#### **5.2.1.1 Data collection and storage**

The data analysis showed that the university offers guidance to researchers dealing with human subjects on the storage and disposal of specific types of research data; those conducting computational research are offered large storage systems. These elements can be recognised as pertinent to research data management and need further consideration, for instance when planning for and developing further services.

In particular, the results showed that researchers are provided with specific information on data storage and disposal, and some consideration is given to the long-term storage of valuable data. Whilst no detailed information on this matter was discussed in the interviews, it would be interesting to investigate whether and how the data from longitudinal or policy-making studies are managed.

However, the data revealed a number of limitations that need further consideration. They are mainly about the responsibility for the data, data storage solutions and disposal of data. Also, it is not clear how the requirements outlined in the “Guidance notes for investigators. Data collection and storage” would be enforced.

Furthermore, one interviewee provided a thorough insight into data practice from the ethical point of view by bringing into consideration such issues like the privacy protection and data publishing/sharing. This clarified that sharing social science data on a larger scale implies new approach to the research design. While data analysis indicated a few issues concerning privacy, such as protection of personal data, it is relatively a new area of concern and research, thus deserves further consideration. It is also worth noting that this was the only example of a formal request to publish the research data in an archive beyond the university.

Finally, the interviewees with IT and computing background and those interested in it discussed the technical aspects of research data management, such as the Institutional Repository and specific services for research computing, like HPC.

### **5.2.2 Stakeholders and roles**

The findings revealed that the provision of research data management is a joint effort involving various stakeholders and roles. The interviewees suggested collectively researchers and academics, the Library, Research Office and IT Services as key stakeholders. They also emphasised that institutional priorities need to be taken into account, and this was supported by the need to engage with schools at Loughborough University and in particular associate deans for research, university managers and senior management when implementing research data management.

Furthermore, the interviewees underlined that researchers, academics were the most important group; some stated with emphasis that their direct engagement was the best way of working. These findings support Loughborough University focus on users which is also reflected in the Library policy.

Other stakeholders mentioned were the High Performance Computing, Research Computing Team, facilities management, Enterprise Office and IP Office that mostly represent technical and business entities of the university.

They are also conveyed in the Working Group; the group is a timely approach to tackle the management of research data at Loughborough University whose participants can contribute different and complementary expertise and insight.

It should also be noted that most of the interviewees are conveners of the group which may affect their views. For instance, two of the interviews were held after the first meeting of the group.

The results imply that there are different entities that have a stake in research data management, and this is supported by the structure and composition of the Working Group. Some of the conveners of this Group, like the Library, Research Office and IT Services represent major support services and play a role in speaking for institutional interests and priorities. Besides, the department of Information Science and the School of Civil and Building Engineering are representatives of researchers and the researcher from the Information Science can offer a theoretical perspective on research data management. On the whole, this is a collective effort that brings together a variety of priorities, interests and responsibilities, therefore, requires further investigation.

#### **4.2.2.1 Prime Responsibility**

The data revealed alternative opinions on which professional group would hold the prime position for taking the lead in research data management.

The contribution of the Library was described as a front end role in research data management, and apparently the academic librarians seemed the key to the overall perception of the Library. On the other hand, a few interviewees recognised the leading role of the Research Office and in part of the Student Research Office, as the departments directly engaged with researchers, especially in the early stages of research practice, and potentially the main reference for research data management. Conversely, the IT Services were perceived as responsible for infrastructure and the back office that could advice colleagues in the library.

Of another opinion were those participants who stated that the library staff was not equipped to, or not focused on supporting research data management and that the IT staff was not tasked to do so. Some interviewees observed that the staff of the Research Office and the Library would be equipped to support RDM with additional training.

Mostly, the provision of research data management was described as a joint effort of the central support services, the Library, Research Office and IT Services, who, as showed previously, were recognised as the key stakeholders.

The results revealed a certain difficulty to indicate who would hold the central role, and this reflects that the management of research data is indeed in the early stages and lacks proven choices. It also indicates that further developments will be guided by local priorities.

#### **4.2.2.2 Collaboration**

The key partners mentioned were industry, national bodies like DCC and other HEIs. This evidence to what extent research data management is a collaborative effort and brings forth other themes of concern that need further consideration, like joint ownership of research data, the role of external organisations and, finally, economies of scale.

The most frequently interviewees suggested the DCC given the partnership between the university and the Centre. Their contribution was described through training sessions, workshops, and also as the provider of reference guides and other support materials for research data management.

In addition, some interviewees mentioned local higher education institutions that were described as potential partners in the RDM and the main reason for this was the ongoing collaboration with local universities and especially among libraries. This was confirmed when in May 2012, according to the High

Performance Computing Midlands webpage, Loughborough University and the University of Leicester with support and funding from the EPSRC initiated a new service that provides cloud supercomputing for academic and industrial partners (HPC Midlands, 2012).

#### **4.2.2.3 Research Councils**

Finally, the last major stakeholder were research councils. The results showed that need to manage research data was presented as twofold: institutional needs of the university and researchers, and the requirements of external agencies, such as research councils. In particular, the interviewees clearly stated or referred to the recent mandate of EPSRC and the importance of this funding agency was described through their contribution to the overall research income of the university.

Discussing the research councils, the mandate to manage research data was described as one of the main reasons for the management of research data; meeting this requirement was presented not only as a mere mandate with which the university needs to comply with but it was also illustrated in a positive and negative light, respectively as an incentive, a lever to improve the current practice and another requirement.

In addition, some interviewees highlighted that the initial requirement, although relevant especially to researchers who apply for bids from EPSRC, will have to satisfy all different kinds of institutional research and would have to be adequate and developed in line with the local needs of researchers and the university.

### **5.2.3 Current issues**

The majority of the interviewees shared quite identical opinions about major issues that arise currently and recognised difficulties facing researchers. Although the RDM was described as still evolving, the interviewees were aware of major aspects and critical factors that hinder the management of research data.

The findings indicated that the majority of the interviewees recognised that research data management is an issue and believed that the most prominent was a lack of central, institutional oversight of research data. In addition, some believed that a central policy was a big gap.

Notably, they recognised drawbacks of the services that they provide and represent; this evidences the institutional focus on enhancement and their attention to users and institutional priorities.

#### **5.2.3.1 Researchers' data practice**

A particular attention was given to the current practice of researchers who, according to most of the interviewees, manage and organise their data inappropriately. Few described researchers as unaware of or not showing interest in the management of research data. Moreover, a few interviewees pointed out that

there was an expressed need of persuading researchers about the value, quality, recognition of their data, and benefits to them.

The findings are partially supported by the documents, namely the ethical guidelines regarding researchers dealing with human subjects, which in part address this potential risk of data loss and were discussed in the *Data collection and storage* section. However, there are no indications of how these requirements would be enforced.

Furthermore, the data practice was described from the institutional perspective as business risk and unavailable information about where data are kept. Conversely, some interviewees observed that lack of proper documentation regarding research data was a potentially helpful knowledge for the university.

The aforementioned considerations emphasize a slight shift of attention from the individual and private practice of researchers towards an institutional viewpoint on the ownership of research data, namely who controls or is responsible for the data. In addition, these considerations indicate that the university is interested to have an oversight of the research data outputs. Although there was no sharp distinction between these two viewpoints, most of the responses showed that researchers are represented as the main responsible for and, by implication, owners of the research data. The university interests in research data holdings can evidence the current attention and concern on this topic.

#### **5.2.3.2 Support services**

Considering the attention that the interviewees pay to the quality of services they provide, the data indicated that support services are not responding to the current needs of researchers. On this regard, the responses indicated that:

- there is no provision of adequate storage for data and interviewees mainly referred to the storage for large scale data and in part to the general data storage,
- training is not specifically on research data management
- research data management is not part of the research lifecycle
- The staff is not equipped or not focused on supporting research data management.

The lack of central coordination can be interpreted as lack of interest, requirement or priority, which was also acknowledged by a few interviewees. In fact, some responses revealed that until recently research data management has not been the priority for the central institution and has become a necessity only with the mandate of major research agencies. However, one interviewee stated that the university has invested in the central storage.

Besides, when discussing how a data archive would be conceived, the interviewees with IT background clearly indicated that the data repository would need to take into account various internal needs for managing data with defined access control.

The above mentioned central storage may or may not be defined as pertaining to research data management, thus another relevant outcome of the data analysis regards the definition of research data management and needs further investigation. One interviewee defined research data management as against ethics and in contrast to data sharing. Therefore, somewhat indirectly, data management was described as data sharing and ethics – imposing restrictions on data. Similarly, other considerations presented data management as parallel to data sharing. Some interviewees observed that re-use of data is a potential not yet considered at the university; they commented the approach of researchers that could be a hurdle to overcome. It should be noted that some of these considerations are a consequence of a personal and professional interest in data sharing.

The results imply that a shared definition of research data management should have been established from the outset of the study. Although no interview question asked directly about a definition of RDM, the responses indicated different concepts that are related to the management of research data. Therefore, the results showed that this concept is still evolving. Such results reinforce the need to further investigate this topic.

#### **5.2.4 Future developments**

The analysis showed different aspects concerning the nearest future and the most significant were further developments of the Institutional Repository, researchers' approach to research data management and aspects that hinder research data management.

It has been noticed that firstly the Working Group would address the issues illustrated in the previous paragraph and that the major effort would focus on the mandate of EPSRC and other agencies. The key elements to address were policy and training.

In particular, some interviewees indicated the difficulty to develop a policy that will fit the priorities of the university and its researchers and, similar considerations were provided about ethical guidelines and other relevant documents that will have to be tailored to the new requirements.

Regarding considerations that provided insight into how a data repository would be conceived, the interviewees with technical background underlined that they would experiment with infrastructure. In addition, even if slightly different opinions were expressed on this regards, namely what types of data

repositories were needed (only metadata or containing different types of data), they showed concern with architecture, access control and security of the infrastructure.

Half of the interviewees expressed contrasting opinions regarding research data management. For instance, two interviewees underlined an overall change of direction that data management would bring, through for instance tailoring guidelines and other institutional documents to the new requirements and a new role that needs new hires. On the other hand, other two interviewees pointed out that it would be a next step and another strand of what they are currently doing.

Other relevant outcomes of the data analysis are the factors concerning research data management that were considered as elements to tackle in the nearest future, namely data selection and appraisal, various data types, data curation and long-term storage, which form elements of the data lifecycle. These findings need further consideration, however, literature showed several examples of pertinent discussion.

Finally, related to the engagement with researchers, a considerable amount of thought was dedicated to finding out how much research data is currently at the university.

#### **5.2.5 Solutions**

The data showed that researchers were recognised as the most important group and were frequently recalled in most of the responses. As a consequence, the interviewees collectively observed that they needed to provide solutions adequate to the needs of researchers and to respond to the problems and day to day activities. Furthermore, it was observed that any new system must be built with users that was considered effective and responsive, and users-researchers were seen as part of the process.

They illustrated the purpose of working closely with researchers by providing specific means of how best do it:

- joint meetings of researchers to ask RDM questions,
- Reach researchers through the word of mouth to spread the knowledge and to be widely known and get involved throughout the research project,
- Organise sessions, workshops, and
- Create blogs, mailing lists.

The data analysis showed that new systems should be as much as possible incremental and employ the existing workflows; it was described as one of the best ways of implementing. The examples were drawn from the activities of the Library, namely the work of academic librarians and the Institutional Repository. In addition, two interviewees pointed out that the current focus of the Library might change and become more research-focused; other responses revealed similar change of directions.



Regarding stakeholders, there are a number of different departments and entities that have a stake in research data management, as showed in the data analysis. Such a partnership may require leadership with clearly defined aims and a shared understanding, and this was clearly expressed in the responses that pointed out the need to have a common definition, good communication and institutional support.

Furthermore, it was highlighted that the Library should not only promote services such as research data management but also shape time for other activities. This was said in regard to the work of academic librarians who were seen as at the forefront of research data management. Without a central structure, they could be provided an advice sheet to distribute to researchers, a small piece of information that would point to further sources of information. It should be noted that such materials are already available, but there are no internal directions or guidelines regarding RDM.

Another significant outcome is about different opinions on a standard approach to implementing research data management. Although the majority of the interviewees highlighted that the materials and tools provided by DCC were recognised as an adequate support and to some extent, a standard, one interviewee pointed out the need to develop a standardised approach.

Finally, from the technical point of view, a variety of solutions were discussed, including economical solutions for the university to provide large storage systems for large datasets and different systems for data repository, ranging from proper repository systems, through content management systems (CMS) to research publications management systems.

## **CHAPTER 6: Conclusions and implications**

In this concluding chapter, the most relevant findings from the data analysis and discussion are aligned with the objectives of the study and are presented in a concise manner. Finally, based on the findings of this research, implications for future research and recommendations for Loughborough University and the University Library are presented.

### **6.1 Research Objectives**

This study aimed to gain an understanding of requirements and opinions of stakeholders involved in the provision of research data management at Loughborough University. Therefore, the following four objectives were outlined to help achieve this aim: 1) Investigate how research data practices are currently supported, 2) Explore how research data management will be implemented, including roles and responsibilities, work practices, 3) Evaluate the outcome of activities taken in the implementation of RDM and 4) Analyse how/whether users are going to be involved in the implementation of RDM.

#### **6.1.1 Examine how research data practices are currently supported**

*The content is deduced from the categories: current activities, stakeholders, issues*

The data analysis showed that Loughborough University is in the early stages in the implementation of research data management. As a consequence, there is no formalised structure nor institutional oversight, policy or coordination, and interviewees stressed this collectively. The results clearly indicated that there are no proven solutions for the management of research data. In fact, undecided and unclear were two concepts reflected in some of the considerations and many interviewees identified details and resulting questions that need further consideration.

However, there are other factors that influence research data management at Loughborough University, namely mandate of research councils, situation in other HEIs, and provision of funding. In brief, the requirements of research councils on research data, similar initiatives in other institutions that started considering this matter, and financial resources that some academic institutions received to support research data management. The results reflect these instances and that indicates that research data management is coming up the agenda in academic institutions.

In order to start considering formally and support actively the management of research data, a Working Group has been set up and comprises a variety of institutional stakeholders, namely the Library, Research Office, IT Services. These can be described as representatives of the central support services. In addition, it conveys two representatives of researchers from the School of Civil and Building Engineering and the department of Information Science. The main purpose of the group is to address the most urgent requirement, namely the mandate of EPSRC and other funding agencies that require researchers and

those that support their work to consider data in the planning for research work and report back on this. Furthermore, the Working Group will address other issues that are discussed in the following paragraphs. As showed in the analysis, the university offers a variety of services and support, ranging from activities of the Research Office to IT Services. Furthermore, most of the interviewees showed a common understanding of responsibilities of the key stakeholders, namely researchers and academics, the Library, Research Office and IT Services. First of all, researchers were described as directly responsible for the research data. Secondly, the Library was perceived as the department that organises and provides access to resources, especially academic librarians who are involved in supporting researchers mainly by means of training and advice. Then, the Research Office was recognised as the main reference for all institutional research that assists research academics throughout their activities and in particular in the early stages of research activity. Finally, IT Services were perceived as responsible for infrastructure and storage, and specific services like High Performance Computing. The evolution of these roles is discussed in the next section.

Undoubtedly, the interviewees presented the current situation from the point of view of background and affiliation giving a rich description from different perspectives. Moreover, the results showed that interviewees mentioned not only the services that they provide, but also other central support services, namely the Research Office, the Library, in particular the work of academic librarians and IT Services. This reciprocal responsiveness and awareness of what others can offer evidence an exemplary collaboration between them; hence there is scope for further collaboration.

In addition, the analysis highlighted a possible involvement of external partners and two main partners were DCC and local higher education institutions. Firstly, the Centre was identified as an authority that would provide training, assistance and a source of materials and tools that support the introduction and development of the management of research data. With regards to local collaboration, a recent project has been initiated by Loughborough University with the university of Leicester and support of ESPRC to provide high performance computing for the academic and private sector. This is an example of economies of scale to tackle the cost of massive amounts of storage for research data. This indicates that external partners like the DCC and research councils may help to implement RDM. In addition, the data showed an increase of research income for Loughborough University. The results imply that research data management is a requirement to sustain the ongoing performance of the university.

Whilst the university has only recently approached the management of research data in a formal way, the interviewees showed a significant awareness of and attention to some of the opportunities and drawbacks of the RDM and recognised that there is a need to provide adequate support services. This self-criticism supports the institutional aim to provide the best user experience and evidences that the institution

constantly strives to enhance.

The data analysis showed such structural drawbacks as until recently the research practice has not been adequately supported, for instance it was observed that, at the moment, the research is currently smaller part of work of academic librarians than their learning support duties. Nevertheless, the Research Office, and especially the Ethical Committee, is now being asked questions about publishing research data and similarly academic librarians, who sometimes are asked about publishing research data.

Similarly, the practice of researchers when dealing with research data was depicted as inappropriate, and the results showed that researchers are managing data individually and provide their own support. Indeed, some authors stressed that researchers are not incorporating appropriate data practices (Humphrey et al., 2000) and chaotic management of digital information (Ogburn, 2010). On the other hand, Doorn & Tjalsma (2007) aptly pointed out that with advances in information technology data services have been democratised making of every researcher a “data supplier” (Doorn & Tjalsma, 2007).

These issues were in the same time considered from the point of view of individual researchers as well as the institution; the university has expressed interest in research outputs that are currently invisible and they have no or little oversight. To sum up the issue of data practice and data holdings has been put as twofold: institutional data and data of researchers. Although the data showed no sharp distinction between these two approaches, it indicates that data property is not clearly defined. Therefore, it deserves further consideration to clarify who owns and who is responsible for the data. Geller drew a similar conclusion regarding intellectual property rights of physical data, like laboratory notebooks (Geller, 2010).

As showed above, interviewees frequently emphasised priorities and interests of the university. However, somewhat surprisingly, little consideration was given to currently adopted institutional documents, policies or operational plans.

In addition, interviewees pointed out a range of issues that need enquiry, such as privacy protection with regards to research data containing personal information, secure storage and disposal of the data. In part, these topics pertain to the evolving structure for research data management at Loughborough University, partially, privacy issues need further consideration.

Final remarks are about the definition of data management. Although one interviewee tried to define the management of data in opposition to ethics that put limits on data sharing, in other considerations interviewees provided some insight into data sharing and open data, thus put data management as related to these concepts. It should be noted that, on the whole, all the interviewees showed interest in preserving, organising and making available research data. These considerations imply that future studies would benefit from considering these aspects jointly.

Interviewees commonly agreed about the gains from research data management, mainly for the institution and the researchers themselves. They observed that research data should be preserved to improve reuse of data, especially to increase profits from the data that are already available. They also emphasised the need to access and validate research outputs. These would be especially beneficial to the university and would enable a central oversight of research outputs.

To sum up:

- RDM is in the early stages and there is no central oversight over this process
- Other HEIs are in a similar situation; some of them received funding to support RDM
- Working Group was set up to approach formally and actively RDM
- Currently there are a number of RDM services and support
- Central support services, like the Library, the Research Office and IT Services have a role in RDM
- External partners, such as DCC and local HEIs have a role in RDM
- There are both institutional issues as well as issues regarding individual researchers and research groups, such as structural drawbacks, data practice of researchers, data property and privacy protection
- Definition of RDM is not clear; it is related to concepts such as data sharing and open data
- The key benefits regarding RDM mainly regard the institution and the researchers

**6.1.2 Explore how RDM support and services will be implemented, including roles and responsibilities, work practices**

*Future and solutions*

The Working Group would design, implement and coordinate RDM processes. In addition, external and industrial partners may have a role in the activities of the Group. This implies that there might be a need for a central structure that would be a reference for research data management at the university.

The Group is now in charge of developing a policy that will address access and re-use of research data and will aim to help and guide researchers. Also, they will set up services, develop procedures and related activities. The data pointed out such priorities as an ongoing support and advice for researchers, technical guidance, training on research data management and data storage. In a similar manner (Delsere, 2008; Hayes et al., 2009) outlined tendencies in the provision of RDM.

The analysis further revealed two main alternatives regarding who would take responsibility for implementing research data management, namely the Library or the Research Office. These two key stakeholders, according to the data, have a similar importance and are perceived as at the forefront of research data management, because of their direct involvement in the research practice. However,

according to the literature (Wilson et al., 2010), this choice depends on local and individual preferences. Finally, most of the responses acknowledged a collaborative provision of research data management by the central support services, namely the Library, the Research Office and IT Services.

Summarising how interviewees perceived the contribution of the key stakeholders to the research community, the Library is focused on organising, providing access to resources, liaison, training, advice and the Institutional Repository; the Research Office is engaged with funders and provides training; the IT Services are responsible for physical infrastructure, represent an indirect reference for research data management and a point of contact and advice for other support services. It is worth noting that these roles are in evolution and reflect the needs of the local community. Whilst the IT Service figure here as a somewhat backend role, literature gives examples when computer faculty and students engaged actively in the research (Williford & Henry, 2012, p. 32). This, in fact, could be the case in any computationally intensive research.

Furthermore, the most significant responses to the issues that are mentioned in the previous section were to engage more directly with researchers, raise awareness and educate them by offering training on research data management and provide adequate tools and structural support for research practice. It has also been stressed to provide support that can address day-to-day, ongoing problems and give solutions, such as a policy on RDM that helps researchers, and tools that support collaborative research. The results indicate that support seemed the first priority making the infrastructure the second priority. Peters & Dryden drew similar conclusion related to large scale research projects (2011, p. 396).

Regarding systems, structure and tools for data repository, the analysis showed different and complementary solutions. The key findings are about systems that are currently associated with institutional repositories and indicate that they would evolve further. The institutional repository and the research publications database were both considered potentially suitable for hosting data outputs; some responses indicated other systems like content management systems (CMS) and commercial solutions for sharing resources.

Considering the main issues that emerged from the data, these are storage and back-up for large amounts of data, access control and a flexible access structure, namely provide adequate infrastructure for storing and backing up research data, define what data will be used, by who and how the access would be provided. In brief, the data identified the following types of repositories for research activity: repository for preservation, for active work and a metadata repository. It is worth noting that the interviewees with technical expertise showed interest in experimenting with a data repository. This was also reflected in another consideration about the universities that should lead the innovation.

Another relevant outcome regards the way that research data management would be developed, and, according to the analysis, this would be achieved by enhancing what is already on offer rather than starting anew. Although, at the moment, there are only a few specific institutional resources on RDM, such as the web pages, the responses indicated no substantial obstacles in enhancing what is currently available. Given that the data indicated the quality and recognition of the current services, these would be adopted to account for RDM, for instance by training staff and increasing research focus of academic librarians. Alternatively, the responses rarely mentioned hires for a new role or a need to change significantly. However, the way forward could be equally distant from these two extremes, as there might be a need to hire for a new role and, contemporary, look how enhance and bend the current support, and this is also apparent from the analysis.

In the end, whilst one interviewee raised the issue about the need to develop a standard approach to RDM, the common thought was that materials provided by the DCC were the main source of support and could be regarded as a standard. Besides, as a few interviewees pointed out, exemplars of implementation of RDM are available and can be looked for in other institutions. Nevertheless, this only confirms that the university needs internal directions for the wider community, and there is an explicit need for support.

To sum up:

- A number of stakeholders are and will be involved in RDM, including the Working Group and some external partners who will shape the process
- One stakeholder may take responsibility for RDM or it may be a joint undertaking
- The key priority regarding RDM will be to develop a policy and support, only then infrastructure
- There are different and complementary solutions to address RDM infrastructure
- The data revealed a common understanding of the main issues that would be addressed
- Interviewees shared a common understanding of the contribution of central support services – Library, Research Office and IT Services
- RDM implementation would be achieved by enhancing current services
- Standard approach to RDM seemed supported by the DCC

### **6.1.3 Evaluate the outcome of activities taken in the implementation of RDM**

The activities at Loughborough University with regard to research data management can best be described as evolving. Yet, even at this early stage, it was possible to evaluate some of the steps taken that have been identified in the data. These include events, activities, roles and responsibilities that have been previously discussed.

Firstly, considering the structure and agenda of the Working Group, it is remarkable that their work has

been initiated by one of the support services, namely the Library and set up together with the Research Office. In addition, it is supported by a representative from IT Services, the department of Information Science and the School of Civil and Building Engineering. It is worth noting the involvement of a researcher from the field of Information Science who can contribute with practical and theoretical.

Literature provides similar examples of initiatives in which the library took a lead (Steinhart et al., 2008; Wilson et al., 2010) and showed their capacity mainly through assistance, support and dedicated services (Peters & Dryden, 2011); a contrasting example, where IT Services had the key responsibility (Wilson et al., 2010). Some authors (Williford & Henry, 2012, p. 33; Wilson et al., 2010) commonly agreed on the importance of institutional collaborations and crossing institutional boundaries.

Contribution of different representatives and their collaboration will ensure a comprehensive endeavour to help introduce and shape the new process, and this can be seen as best practice, given that a range of stakeholders can bring different but complementary expertise and point of views. Besides, the collaboration between the key stakeholders is significant for the development of strategic plans of the university and this was also aptly noticed by one interviewee. Furthermore, some interviewees observed that the Group needs a strong leadership, communication and a shared understanding. The findings indicate that leaderships and communication would contribute to the progress and success of the activities, which is clear if they aim to develop a strategic programme and proceed with further plans. Such joint forces revealed another strategic aspect that resulted from the analysis, a good reciprocal understanding and internal collaboration within the library; this was pointed out by Hswe & Holt (2011, p. 13).

In the second place, from the outset of the Working Group, interviewees stressed that researchers need to be part of the process, and this is supported by the literature (Martinez-Urbe & Macdonald, 2009). In fact, they are already two research representatives in the Working Group and, according to the results, further on their number could increase. It is worth noting that this involvement of researchers, although broadly stressed in the literature, is not always reflected in practice.

The Group shall address the requirement of EPSRC and other main research agencies to manage research data; however, as several interviewees aptly pointed out, their programme is developing gradually. The university will develop a policy about RDM and this resembles what other academic institutions are doing (Hswe & Holt, 2011). Further similarities can be observed, as the study showed that several services are already provided and they have no single oversight; the same findings were demonstrated in (Peters & Dryden, 2011, p. 397). Even if a number of persisting questions regarding centralised services was evidenced, these doubts were also exposed in the literature (Hswe & Holt, 2011, p. 16)



Given that research data management would be financed internally, therefore, it is not surprising to note that most of the interviewees stressed the importance of having support from the university in terms of human, financial resources and recognition of senior management.

Moreover, as aforementioned, the university is a Framework University for EPSRC, which, according to the agreement between the funder and the institution (EPSRC, 2012), implies that they provide an example to and share best practice with other institutions. In addition, the university has obtained support of the DCC to prepare the staff to deal with the RDM independently. This confirms the importance and role of external and national bodies that can provide guidance, resources and support to institutions. In addition, DCC is an example of support in the form of electronic resources that are available to all institutions nationally and worldwide. On the whole, it can be deduced that even without a financial support from outside, institutional and professional commitment can prove effective and should be recognised.

From the analysis, it is also clear that the interviewees expressed attention towards a range of requirements on various levels, like the needs of researchers, the institutional needs and those of stakeholders. The issues regarding researchers are further discussed in the next section. This shows that the local priorities are most highly valued and have to be outlined, understood and taken into account. All in all, this is a joined initiative that has come up to respond to the requirement of funding councils, as a shared understanding of issues and was based on a small scale study of research requirements.

Several of these considerations regard sharing, collaboration and international effort, which are increasingly common to libraries. In many instances, these elements seem essential to the implementation of RDM and, for instance, can enable economies of scale through collaboration between local higher education institutions to address some demanding issues, like storage.

On top of that, supporting examples can be found in the literature showing developments of initiatives to address crucial issues that are based on available resources, including materials and tools developed by other institutions and using staff and procedures that are already in place (Peters & Dryden, 2011; Thomas, 2011).

#### To sum up:

- RDM Working Group has been initiated by the Library and set up together with the Research Office; there is also a representative from IT Services, one academic school and one department
- Contribution of different stakeholders and representatives will help introduce and shape the new process, and this can be seen as best practice, given that a range of stakeholders can bring different but complementary expertise and point of views

- Two representatives of researchers are sitting on the Working Group and more could be involved
- The Group shall address the requirement of EPSRC and other main research agencies to manage research data
- Institutional support to RDM will be fundamental to help achieve a successful implementation
- According to the Framework Agreement with EPSRC Loughborough University will share best practice
- A range of requirements on various levels, like the needs of researchers, the institutional needs and those of stakeholders will be taken into consideration

#### **6.1.4 Analyse whether/how users are going to be involved in the implementation of RDM**

*N.B. researchers are considered users in this study*

The data analysis showed awareness and attention of the interviewees towards researchers who described engagement of researchers as one of the most fundamental aspects of implementing research data management. Such a recognition of the importance and direct engagement of researchers are a best practice that support the Loughborough University focus on users (Loughborough University, 2007). Also, some interviewees described their own experience and background in research, which could help them better understand the issues that researchers are facing.

Interviewees not only reflected upon theoretically how to engage directly with users, but also provided several examples and different factors concerning users' involvement, such as receptivity of researchers, their approach to data management and to data sharing.

Discussing a general approach to implementing RDM, the analysis indicated what could be called as not intrusive means of collaboration, namely interviewees suggested avoiding dictating from on high, inflicting and discussing together. These are examples of best practice and can prove effective when introducing new processes. Collectively, as indicated above, the interviewees mentioned working closely with researchers and especially sitting on the research activity and several authors broadly support this with real life examples (Bracke, 2011; Delserone, 2008; Freiman, Ward, & Jones, 2010).

In addition, interviewees discussed making contacts with researchers and acknowledged that collaboration of different stakeholders was essential to succeed. For this reason, the Research Office was recognised as the main reference for research staff and other institutional stakeholders. They were considered equally indispensable to both because of their close relation with researchers.

Finally, the analysis showed several means of engaging with researchers and all deserved further consideration. They were:

- Share feedback of the Working Group representatives with the research community
- Organise drop in sessions, brown bag lunches during which researchers could ask questions
- Join meetings of research staff to ask questions about research data management
- Set up a mailing list of research staff to facilitate collaboration
- Initiate working with a small group of researchers interested in data sharing
- Initiate collaboration with a designated community, such as computing research
- Facilitate and offer possibilities for research collaboration between researchers from local HEIs; this is similar to (Ogburn, 2010, p. 244)
- Provide not excessive amount of information in the point of need, such as an information sheet, might be also through a webpage
- Communicate with researchers through research blog
- Use word of mouth to spread the news efficiently (this is based on personal contacts between researchers and those supporting them, is highly subjective)
- Show examples of data assessment and value
- Indicate examples of risk associated with lack of data management
- Recognise researchers as the most crucial for data appraisal and involve them in data management

The main conceptual issues regarding researchers were lack of recognition and assessment of the research data and that some researchers do not envisage the role of the Library in their practice. Similar considerations regarding other stakeholders were not provided. It is not clear from the data whether this is specific to the Library, however, the literature supports this showing that some researchers do not seek library's assistance (Meyer et al., 2011; Peters & Dryden, 2011). Therefore, further consideration is needed to clarify the solutions for collaboration that could prove difficult.

On this regard Lage et al. suggested to partner with researchers who were receptive to a library role in data curation. Notably, they discussed several factors that influence this disposition of researchers in terms of positive and negative correlations to the receptivity of library involvement (Lage et al., 2011, pp. 931–932). Apparently, it seems that similar issues are relevant to any collaboration and are not specific to RDM. The remaining issue is how to persuade researchers and raise awareness about the value of the research data. That is to say, how and to what extent the institution can influence receptivity of researchers and this also deserves further inquiry. Likewise, other issues are worth considering, such as how much control academic institutions have over research outputs and what are models and frameworks that can clarify it.

To sum up:

- Interviewees recognised the importance and direct engagement of researchers as the most fundamental aspect of implementing RDM
- Interviewees provided several examples and different factors concerning users' involvement
- Working closely with researchers and sitting on the research activity were mentioned to engage with researchers
- A general approach to RDM was to avoid dictating from on high, inflicting and to discuss together
- Conceptual issues regarding researchers were lack of recognition and assessment of the research data and lack of recognition of the role of the Library

## **6.2 Recommendations for Loughborough University and the University Library**

Based on the data analysis, there are a few points that could be considered and may contribute to the continuing success of the process:

- Estimate data management provisions within support services, determine the current structure for technology support and identify the general configuration of networks, data storage capacities provided and technological support.
- Identify policies and institutional documents that contain elements of RDM and try to adapt them
- Identify a leading department for RDM and main researcher support “officers”
- Identify a shared definition of research data management and a message to deliver to researchers and within the university, clarify the property of research data
- (Re)consider how much time and resources the central support service dedicate to RDM
- Develop procedures and related activities, clarify front and back office roles
- Provide support to staff, by developing or pointing to materials, tools (reuse)
- Link up scattered activities and roles, by creating one reference point on RDM
- Identify schools, departments or research groups that are interested in data sharing and those that are against it
- Estimate the volume of research data holdings, identify managed and unmanaged research data, i.e. data from longitudinal or policy-making studies
- Identify requirements for storage and back-up and provide guidelines

## **6.3 Implications for further research**

The results indicated that there are no proven solutions for the management of research data and this was reflected in some considerations of stakeholders who identified details and resulting questions that need further enquiry. Therefore, further studies should consider a few points:

- Identify definition of research data management from the point of view of practitioners in academic institutions
- Explore approaches of librarians towards RDM
- explore and compare RDM workflows in libraries
- Identify standards on RDM and who contributes to developing standards
- Explore and compare examples of implementation of research data management, in particular support/guidance and systems
- Explore and analyse systems, or data repositories, their specifications and features
- Explore examples of local collaboration to provide RDM
- Identify issues regarding the management of jointly owned data
- Identify best practices of managing research data
- Explore jointly research data management and open data, data sharing and other relevant aspects
- Identify and compare automated metadata generation for research data and manual methods

## References

- Accart, J.-P. (2011). The Open Science: the OAI7 Workshop in Geneva (Switzerland) – June 2011. *Library Hi Tech News*, 28(7), 1–4. doi:10.1108/07419051111184016
- Akmon, D., Zimmerman, A., Daniels, M., & Hedstrom, M. (2011). The application of archival concepts to a data-intensive environment: working with scientists to understand data management and preservation needs. *Archival Science*, 11(3-4), 329–348. doi:10.1007/s10502-011-9151-4
- Alvaro, E., Brooks, H., Ham, M., Poegel, S., & Rosencrans, S. (2011). E-Science Librarianship: Field Undefined. *Issues in Science and Technology Librarianship*. doi:10.5062/F46Q1V55
- Angevaare, I. (2009). Taking Care of Digital Collections and Data: “Curation” and Organisational Choices for Research Libraries. *Liber Quarterly: The Journal of European Research Libraries*, 19(1), 1–12. doi:Article
- ARL/NSF Workshop. (2006). *To Stand the Test of Time: Long-term Stewardship of Digital Data Sets in Science and Engineering*. Association of Research Libraries (ARL). Retrieved from <http://www.arl.org/pp/access/nsfworkshop.shtml>
- Ashby, M., Espinet, E. O., Fry, J., Lund, P., Stubbings, R., & Walton, G. (2011). *Scholarly communication at Loughborough University 2011*. Loughborough University Library. Retrieved from <http://www.lboro.ac.uk/library/about/PDFs/Scholarly%20communication%20at%20Loughborough%20University%202011.pdf>
- Association of Research Libraries. (2004, October). E-Research and Supporting Cyberinfrastructure. *Association of Research Libraries*. Retrieved February 29, 2012, from <http://www.arl.org/resources/pubs/fallforumproceedings/forum04proceedings.shtml>
- Auckland, M. (2012). *Re-skilling for Research*. Research Libraries UK. Retrieved from <http://www.rluk.ac.uk/content/re-skilling-research>
- Bampton, R., & Cowton, C. J. (2002). The E-Interview. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 3(2). Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/848>
- Barwick, J. (2007). Building an institutional repository at Loughborough University: some experiences.

*Program: electronic library and information systems*, 41(2), 113–123.

doi:10.1108/00330330710742890

Beagrie, N. (2006). Digital Curation for Science, Digital Libraries, and Individuals. *International Journal of Digital Curation*, 1(1), 3–16. doi:10.2218/ijdc.v1i1.2

Beagrie, N., Beagrie, R., & Rowlands, I. (2009). Research data preservation and access: the views of researchers'. *Ariadne*, (60). Retrieved from <http://www.ariadne.ac.uk/issue60/beagrie-et-al/>

Berlin Declaration. (2003). Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Retrieved from <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung/>

Berman, F. (2008). Got data?: a guide to data preservation in the information age. *Communications of the ACM*, 51(12), 50–56.

Bracke, M. S. (2011). Emerging Data Curation Roles for Librarians: A Case Study of Agricultural Data. *Journal of Agricultural & Food Information*, 12(1), 65–74. doi:10.1080/10496505.2011.539158

Brandt, D. S. (2007). Librarians as partners in e-research: Purdue university libraries promote collaboration. *College & Research Libraries News*, 68(6). Retrieved from <http://crln.acrl.org/content/68/6/365.full.pdf>

Brophy, P. (2007). Communicating the library: librarians and faculty in dialogue. *Library Management*, 28(8/9), 515–523. doi:10.1108/01435120710837792

Candela, L., Castelli, D., & Pagano, P. (2009). On-demand virtual research environments and the changing roles of librarians. *Library Hi Tech*, 27(2), 239–251.

doi:10.1108/07378830910968191

Caplan, P. (2012). On discovery tools, OPACs and the motion of library language. *Library Hi Tech*, 30(1), 108–115. doi:10.1108/07378831211213247

Carlson, Jake. (2012). Demystifying the data interview: Developing a foundation for reference librarians to talk with researchers about their data. *Reference Services Review*, 40(1), 7–23.

doi:10.1108/00907321211203603

Carlson, Jacob, Fosmire, M., Miller, C. C., & Nelson, M. S. (2011). Determining Data Information Literacy Needs: A Study of Students and Research Faculty. *portal: Libraries & the Academy*, 11(2),

629–657.

Carlson, S. (2006). Lost in a Sea of Science Data. *The Chronicle of Higher Education*, 52(42), A35–A36.

Choudhury, G. S. (2008). Case Study in Data Curation at Johns Hopkins University. *Library Trends*, 57(2), 211–220. doi:10.1353/lib.0.0028

Cooper, R. (2009). Online interviewing: It's not as simple as point and click. *The Weekly Qualitative Report*, 2(43), 250–253.

Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2010). Data Sharing, Small Science and Institutional Repositories. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1926), 4023–4038. doi:10.1098/rsta.2010.0165

Data, data everywhere. (2010, February 25). *The Economist*. Retrieved from

<http://www.economist.com/node/15557443>

de Cock Buning, M., Ringnalda, A., & van der Linden, T. (2009). *The legal status of raw data: a guide for research practice*. Utrecht: SURF Foundation. Retrieved from

[http://www.surf.nl/en/publicaties/Documents/SURFdirect\\_De%20juridische%20status%20van%20ruwe%20data\\_wegwijzer\\_ENG.pdf](http://www.surf.nl/en/publicaties/Documents/SURFdirect_De%20juridische%20status%20van%20ruwe%20data_wegwijzer_ENG.pdf)

Delserone, L. M. (2008). At the Watershed: Preparing for Research Data Management and Stewardship at the University of Minnesota Libraries. *Library Trends*, 57(2), 202–210.

Denison, T., Kethers, S., & McPhee, N. (2007). Managing the soft issues in e-research: a role for libraries? *Australian Academic & Research Libraries*, 38(1), 1–14.

Denscombe, M. (2003). *The Good Research Guide: For Small-scale Social Research Projects* (2nd ed.). Open University Press.

Denscombe, M. (2009). *Ground Rules for Social Research: Guidelines for Good Practice*. McGraw-Hill.

Digital Curation Centre. (2012). What is digital curation? Retrieved June 16, 2012, from

<http://www.dcc.ac.uk/digital-curation/what-digital-curation>

Doorn, P., & Tjalsma, H. (2007). Introduction: archiving research data. *Archival Science*, 7(1), 1–20.

doi:10.1007/s10502-007-9054-6

Dorner, D. G., & Revell, J. T. (2011). Subject Librarians' Perceptions of Institutional Repositories as an



- Information Resource. *Online Information Review*, 36(2), 6–6.
- Engineering and Physical Science Research Council, E. (2011, July 8). Principles. Retrieved June 26, 2012, from <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/principles.aspx>
- Engineering and Physical Science Research Council, E. (2012, June 21). EPSRC Policy Framework on Research Data. Retrieved June 26, 2012, from <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/default.aspx>
- EPSRC. (2012). University framework agreements. Retrieved June 27, 2012, from <http://www.epsrc.ac.uk/about/partner/universities/Pages/frameworkagreements.aspx>
- eVALUED. (2006). Interviews. *Evalued: An evaluation toolkit for e-library developments*. Retrieved March 7, 2012, from <http://www.evalued.bcu.ac.uk/tutorial/4c.htm>
- Finance Office. (2012). Financial Statements 2010-2011. *Loughborough University*. Retrieved June 26, 2012, from <http://www.lboro.ac.uk/admin/financial/pages/statements.htm>
- Fowler, S. (2012). *Managing Research Data: a pilot study in Health and Life Sciences: The UWE Case Study*. Bristol: University of the West of England. Retrieved from <http://www1.uwe.ac.uk/library/usingthelibrary/servicesforresearchers/datamanagement/managingresearchdata/projectoutputs/workpackages1and2.aspx>
- Freiman, L., Ward, C., Jones, S., Molloy, L., & Snow, K. (2010). *Incremental: scoping study and implementation plan*. Cambridge, Oxford: University of Cambridge, University of Glasgow. Retrieved from [http://www.lib.cam.ac.uk/preservation/incremental/Incremental\\_Scoping\\_Report\\_062010.pdf](http://www.lib.cam.ac.uk/preservation/incremental/Incremental_Scoping_Report_062010.pdf)
- Gabridge, T. (2009). The Last Mile: Liaison Roles in Curating Science and Engineering Research Data. *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC*, 265, 15–21.
- Geller, L. (2010). Data Management in Academic Settings: An Intellectual Property Perspective. *Science and Engineering Ethics*, 16(4), 769–775. doi:10.1007/s11948-010-9240-4
- Genoni, P., Merrick, H., & Willson, M. A. (2006). Scholarly communities, e-research literacy and the academic librarian. *Electronic Library, The*, 24(6), 734–746. doi:10.1108/02640470610714189
- Gerring, J. (2007). *Case study research*. Cambridge University Press. Retrieved from

<http://www.dawsonera.com/depp/reader/protected/external/EBookView/S9780511319686/S15>

Gibson, L. (2010, June). Using Email Interviews. The University of Manchester. Retrieved from <http://www.socialsciences.manchester.ac.uk/morgancentre/realities/toolkits/email-interviews/index.html>

Gillham, B. (2005). *Research Interviewing: The Range of Techniques*. Berkshire, , GBR: McGraw-Hill Education. Retrieved from <http://site.ebrary.com/lib/tallinn/docDetail.action?docID=10161349>

Gold, A. (2007a). Cyberinfrastructure, Data, and Libraries, Part 2: Libraries and the Data Challenge: Roles and Actions for Libraries. *D-Lib Magazine*, 13(9/10). doi:10.1045/july20september-gold-pt2

Gold, A. (2007b). Cyberinfrastructure, Data, and Libraries, Part 1: A Cyberinfrastructure Primer for Librarians. *D-Lib Magazine*, 13(9/10). doi:10.1045/september20september-gold-pt1  
Cyberinfrastructure, Data, and Libraries, Part 1: A Cyberinfrastruct..  
<http://www.dlib.org/dlib/september07/gold/09gold-pt1.html> 12

Gold, A. (2010). Data Curation and Libraries: Short-Term Developments, Long-Term Prospects. *Office of the Dean (Library)*. Retrieved from <http://works.bepress.com/agold01/9/>

Goldenberg-Hart, D. (2004). Libraries and Changing Research Practices: A Report of the ARL/CNI Forum on E-Research and Cyberinfrastructure. *ARL*, 237. Retrieved from <http://hdl.handle.net/10150/105726>

Gorman, G. E., Clayton, P., Shep, S. J., & Clayton, A. (2005). *Qualitative research for the information professional: a practical handbook* (2nd ed.). Facet Publishing. Retrieved from <http://www.dawsonera.com/depp/reader/protected/external/EBookView/S9781856047982/S24>

Greifeneder, E. (2010). A content analysis on the use of methods in online user research. *Digital Library Future: User perspectives and institutional strategies*. Munich: K.G.Saur.

Groenewald, R., & Breytenbach, A. (2011). The use of metadata and preservation methods for

continuous access to digital data. *Electronic Library, The*, 29(2), 236–248.

doi:10.1108/02640471111125195

Haas, J., & Murphy, S. (2009). E-Science and Libraries: Finding the Right Path. *Issues in Science and Technology Librarianship, Viewpoints*, 57. Retrieved from

<http://www.istl.org/09-spring/viewpoint1.html>

HATII, & University of Glasgow. (2009, October). Data asset framework: implementation guide.

Retrieved from [http://www.data-audit.eu/docs/DAF\\_Implementation\\_Guide.pdf](http://www.data-audit.eu/docs/DAF_Implementation_Guide.pdf)

Hayes, B., Harroun, J., & Temple, B. (2009). Data management and curation of research data in academic scientific research environments. *Proceedings of the American Society for Information Science and Technology*, 46(1), 1–4. doi:10.1002/meet.2009.1450460368

Heidorn, P. B. (2008). Shedding Light on the Dark Data in the Long Tail of Science. *Library Trends*, 57(2), 280–299. doi:10.1353/lib.0.0036

Heidorn, P. B. (2011). The Emerging Role of Libraries in Data Curation and E-science. *Journal of Library Administration*, 51(7-8), 662–672. doi:10.1080/01930826.2011.601269

Henty, M., Weaver, B., Bradbury, S., & Porter, S. (2008). *Investigating Data Management Practices in Australian Universities*. Australian Partnership for Sustainable Repositories (APSR). Retrieved from [http://www.apsr.edu.au/investigating\\_data\\_management](http://www.apsr.edu.au/investigating_data_management)

Hey, T., & Hey, J. (2006). e-Science and its implications for the library community. *Library Hi Tech*, 24(4), 515–528. doi:10.1108/07378830610715383

Hey, T., & Trefethen, A. (2003). The Data Deluge: An e-Science Perspective. *Grid Computing - Making the Global Infrastructure a Reality*, 809–824.

Hockx-Yu, H. (2006). Digital preservation in the context of institutional repositories. *Program: electronic library and information systems*, 40(3), 232–243. doi:10.1108/00330330610681312

HPC Midlands. (2012). HPC Midlands Cloud Supercomputing for Academia and Industry. Retrieved June 27, 2012, from <http://hpc-midlands.ac.uk/>

Hswe, P., & Holt, A. (2011). *Joining in the enterprise of response in the wake of the NSF data management planning requirement* ( No. 274). Research Library Issues: A Bimonthly Report from ARL, CNI,

- and SPARC (pp. 11–17). Retrieved from <http://publications.arl.org/rli274/12>
- Humphrey, C. K., Estabrooks, C. A., Norris, J. R., Smith, J. E., & Hesketh, K. L. (2000). Archivist on Board: Contributions to the Research Team. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(3). Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/1022/2203>
- James, N., & Busher, H. (2009). *Online Interviewing*. SAGE. Retrieved from <http://www.uk.sagepub.com/books/Book231519?prodId=Book231519&>
- Joint, N. (2007). Data preservation, the new science and the practitioner librarian. *Library Review*, 56(6), 451–455. doi:10.1108/00242530710760337
- Kallenborn, R., & Becker, C. (2009). Digital discovery: strategies and solutions: Report on the 29th annual Conference of the International Association of Technological University Libraries (IATUL) held in Auckland, New Zealand, 21-24 April 2008. *New Library World*, 110(5/6), 280–290. doi:10.1108/03074800910954299
- Kvale, S. (1996). *InterViews: An Introduction to Qualitative Research Interviewing* (illustrated ed.). Sage Publications, Inc.
- Lage, K., Losoff, B., & Maness, J. (2011). Receptivity to Library Involvement in Scientific Data Curation: A Case Study at the University of Colorado Boulder. *portal: Libraries and the Academy*, 11(4), 915–937. doi:10.1353/pla.2011.0049
- Lankes, R. D., Cogburn, D., Oakleaf, M., & Stanton, J. (2008). Cyberinfrastructure Facilitators: New Approaches to Information Professionals for E-Research. Retrieved Oct, 25(September).
- Lewis, M. (2008). IC and e-SCI: Postcards from the Egg. *New Review of Academic Librarianship*, 14(1-2), 36–54. doi:10.1080/13614530802519121
- Lord, P., & Macdonald, A. (2003). *Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision*. The JISC Committee for the Support of Research.
- Loughborough University. (2007). *Loughborough University – Towards 2016 Strategic Plan – 2006/07 edition*. Loughborough University. Retrieved from [http://www.lboro.ac.uk/admin/policy/strategic\\_plan.pdf](http://www.lboro.ac.uk/admin/policy/strategic_plan.pdf)

- Loughborough University. (2012a). About the University: Our achievements. Retrieved June 26, 2012, from <http://www.lboro.ac.uk/about/achievements/index.html>
- Loughborough University. (2012b). About the University: research. *Loughborough University*. Retrieved June 26, 2012, from <http://www.lboro.ac.uk/about/research/index.html>
- Loughborough University. (2012c). Open Data projects at Loughborough University. Retrieved June 27, 2012, from <http://open.lboro.ac.uk/credits/>
- Loughborough University Library. (2011). *Loughborough University Library strategic plan 2010-2013*. Loughborough University. Retrieved from <http://www.lboro.ac.uk/library/about/PDFs/strapl-10-13.pdf>
- Lynch, C., & Carleton, D. E. (2009). Lecture: Impact of Digital Scholarship on Research Libraries. *Journal of Library Administration*, 49(3), 227–244. doi:10.1080/01930820902785041
- Macdonald, A., & Lord, P. (2003). *Digital Data Curation Task Force: report of the Task Force Strategy Discussion Day*. JISC. Retrieved from [http://www.jisc.ac.uk/uploaded\\_documents/CurationTaskForceFinal1.pdf](http://www.jisc.ac.uk/uploaded_documents/CurationTaskForceFinal1.pdf)
- Martinez-Urbe, L. (2009). *Research data management services: findings of the consultation with service providers* (Text). Scoping Digital Repository Services for Research Data Management - Research Data Management Services. University of Oxford. Retrieved from <http://www.ict.ox.ac.uk/odit/projects/digitalrepository/docs/ConsultationWithOxfordServiceProviders.pdf>
- Martinez-Urbe, L., & Macdonald, S. (2009). User Engagement in Research Data Curation. In M. Agosti, J. Borbinha, S. Kapidakis, C. Papatheodorou, & G. Tsakonas (Eds.), *Research and Advanced Technology for Digital Libraries* (Vol. 5714, pp. 309–314). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from <http://www.springerlink.com/content/7mnq13x34717p483>
- McKay, D. (2010). Oranges Are Not the Only Fruit: An Institutional Case Study Demonstrating Why Data Digital Libraries Are Not the Whole Answer to E-Research. In G. Chowdhury, C. Koo, & J. Hunter (Eds.), *The Role of Digital Libraries in a Time of Global Change* (Vol. 6102, pp. 236–249). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from

- <http://www.springerlink.com.ezproxy.tlu.ee/content/w25462w2qqt14738/>
- Meyer, E. T., Bulger, M., Kyriakidou-Zacharoudiou, A., Power, L., Williams, P., Venters, W., Terras, M., et al. (2011). *Collaborative yet independent: Information practices in the physical sciences*. Research Information Network. Retrieved from [www.rin.ac.uk/phys-sci-case](http://www.rin.ac.uk/phys-sci-case)
- Newton, M. P., Miller, C. C., & Bracke, M. S. (2010). Librarian Roles in Institutional Repository Data Set Collecting: Outcomes of a Research Library Task Force. *Collection Management*, 36(1), 53–67. doi:10.1080/01462679.2011.530546
- OECD. (2007). *Principles and Guidelines for Access to Research Data from Public Funding*. OECD.
- Ogburn, J. L. (2010). The Imperative for Data Curation. *portal: Libraries and the Academy*, 10(2), 241–246. doi:10.1353/pla.0.0100
- Osswald, A. (2008). E-science and information services: a missing link in the context of digital libraries. *Online Information Review*, 32(4), 516–523. doi:10.1108/14684520810897395
- Paterson, M., Lindsay, D., Monotti, A., & Chin, A. (2007). DART: a new missile in Australia's e-research strategy. *Online Information Review*, 31(2), 116–134. doi:10.1108/14684520710747185
- Peräkylä, A., & Ruusuvuori, J. (2011). Analyzing talk and text. *The SAGE Handbook of Qualitative Research* (Fourth.). Lon: SAGE. Retrieved from <https://helda.helsinki.fi/handle/10138/29485>
- Peters, C., & Dryden, A. R. (2011). Assessing the Academic Library's Role in Campus-Wide Research Data Management: A First Step at the University of Houston. *Science & Technology Libraries*, 30(4), 387–403. doi:10.1080/0194262X.2011.626340
- Pickard, A. J. (2007). *Research methods in information*. London: Facet Publishing.
- Planning Team. (2012). Facts and Figures. *Loughborough University*. Retrieved January 27, 2012, from <http://www.lboro.ac.uk/admin/ar/planning/facts.htm>
- Research Councils UK. (2012). Common Principles on Data Policy. Retrieved June 26, 2012, from <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>
- Research Support. (2011). *Operational plan 2011/2012*. Loughborough University Library.
- Rusbridge, C., Burnhill, P., Ross, S., Buneman, P., Giaretta, D., Lyon, L., & Atkinson, M. (2005). The Digital Curation Centre: A Vision for Digital Curation. *Proceedings From Local to Global: Data*

- Interoperability--Challenges and Technologies* (pp. 1–11). Sardinia, Italy. Retrieved from <http://eprints.erpanet.org/82/>
- Sade-Beck, L. (2004). Internet ethnography: Online and offline. *International Journal of Qualitative Methods*, 3(2), 1–14.
- Salo, D. (2007). Innkeeper at the Roach Motel. Retrieved from <http://digital.library.wisc.edu/1793/22088>
- Science Staff. (2011). Challenges and Opportunities. *Science*, Special Online Collection: Dealing with Data, 331(6018). doi:10.1126/science.331.6018.692
- Simpson, P., & Hey, J. (2006). Repositories for research: Southampton's evolving role in the knowledge cycle. *Program: electronic library and information systems*, 40(3), 224–231. doi:10.1108/00330330610681303
- Soehner, C., Steeves, C., & Ward, J. (2010). *E-Science and Data Support Services: A Study of ARL Member Institutions*. Washington, DC: Association of Research Libraries. Retrieved from [http://www.arl.org/bm~doc/escience\\_report2010.pdf](http://www.arl.org/bm~doc/escience_report2010.pdf)
- Steinhart, G., Saylor, J., Albert, P., Alpi, K., Baxter, P., Brown, E., Chiang, K., et al. (2008). *Digital Research Data Curation: Overview of Issues, Current Activities, and Opportunities for the Cornell University Library*. Cornell University Library. Retrieved from <http://hdl.handle.net/1813/10903>
- Swan, A., & Brown, S. (2008). *Skills, role and career structure of data scientists and curators: an assessment of current practice and future needs*. JISC, Key Perspectives. Retrieved from <http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareersfinalreport.pdf>
- Tellis, W. (1997). Introduction to Case Study. *The Qualitative Report*, 3(2). Retrieved from <http://www.nova.edu/ssss/QR/QR3-2/tellis1.html>
- Thomas, J. (2011). Future-proofing: the academic library's role in e-research support. *Library Management*, 32(1/2), 37–47. doi:10.1108/01435121111102566
- Trochim, W. M. K. (2006). Qualitative Measures. *Research methods: knowledge base*. Retrieved March 8, 2012, from <http://www.socialresearchmethods.net/kb/qual.php>

- University Library. (2012). Facts and figures. *Loughborough University*. Retrieved June 23, 2012, from <http://www.lboro.ac.uk/services/library/about/facts/>
- Van de Sompel, H., Payette, S., Erickson, J., Lagoze, C., & Warner, S. (2004). Rethinking Scholarly Communication. *D-Lib Magazine*, 10(9). doi:10.1045/september2004-vandesompel
- Van den Eynden, V., Corti, L., Woollard, M., Bishop, L., & Horton, L. (2011). *Managing and sharing data: best practices for researchers*. University of Essex: UK Data Archive. Retrieved from <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>
- Williford, C., & Henry, C. (2012). *One Culture. Computationally Intensive Research in the Humanities and Social Sciences: a report on the experiences of first respondents to the Digging into Data Challenge*. Washington, D.C.: Council on Library and Information Resources. Retrieved from <http://www.clir.org/pubs/reports/pub151>
- Wilson, J. A. J., Fraser, M. A., Martinez-Urbe, L., Patrick, M., Akram, A., & Mansoori, T. (2010). Developing Infrastructure for Research Data Management at the University of Oxford. *Ariadne*, (65). Retrieved from <http://www.ariadne.ac.uk/issue65/wilson-et-al/>
- Witt, M., Carlson, J., Brandt, D. S., & Cragin, M. H. (2009). Constructing Data Curation Profiles. *International Journal of Digital Curation*, 4(3), 93–103. doi:10.2218/ijdc.v4i3.117
- Yakel, E. (2007). Digital curation. *OCLC Systems & Services*, 23(4), 335–340. doi:10.1108/10650750710831466
- Zhao, Y. (2009). Changing of library services under e-research environment. *Electronic Library, The*, 27(2), 342–348. doi:10.1108/02640470910947683



# Appendices

## Appendix 1: Introductory email

Dear XX

subject: invitation to participate in dissertation study on research data management

As I have mentioned to you during my internship at the Loughborough University Library, I am a student of International Master in Digital Library Learning (DILL) at the University of Parma undertaking a study on “**Research data management: a case at the Loughborough University Library**”.

The purpose of this study is to explore how the Loughborough University Library is embracing research data support services. More specifically it will look at requirements, roles and responsibilities of some of the stakeholders in the implementation of research data management. Therefore I would like to find answers to such questions as:

- how research and data practices are currently supported
- how research data management support and services will be implemented, including roles and responsibilities
- what are the requirements of some of the stakeholders involved
- how/whether the users are going to be involved

It is hoped that the results from this research will benefit Loughborough University through clarifying requirements for services to manage and curate research data. This research study forms a major component of the requirements for the fulfilment of the masters in DILL.

In order to elicit the required information I would like to invite you to take part in a **half-hour online interview**. If you are able to take part, please could you advise me as to when you might be available? I am happy to meet you in March at any time or please indicate a date that suits you best.

Best regards

Ewelina Melnarowicz

## Appendix 2: Participant Information Sheet

### **Research data management: a case at the Loughborough University Library** **Participant Information Sheet**

Ewelina Melnarowicz, International Master in Digital Library Learning (DILL),  
emelnarowicz@gmail.com  
Pat Dixon, Supervisor, dixonap@aol.com

#### **What is the purpose of the study?**

The purpose of this study is to explore how the Loughborough University Library is embracing research data support services. More specifically it will look at requirements, roles and responsibilities of some of the stakeholders in the implementation of research data management. It is hoped that the results from this research will benefit Loughborough University through clarifying requirements for services to manage and curate research data. The research questions are:

- how research and data practices are currently supported
- how RDM support and services will be implemented, including roles and responsibilities
- how/whether the users are going to be involved in the implementation

#### **Who is doing this research and why?**

This research study forms a major component of the requirements for the fulfilment of the masters in DILL.

#### **Once I take part, can I change my mind?**

You can withdraw at any time, for any reason and you will not be asked to explain your reasons for withdrawing.

#### **Will I be required to attend any sessions and where will these be?**

Participation in an half-hour online interview. With your permission I would like to record this interview. The recording will be used for the purpose of this research only and the transcripts and recordings will be kept until the end of the project, June 2012.

#### **What personal information will be required from me?**

All information that you provide will be treated with confidentiality and will be used only after your signed permission and your amendments of the interview transcripts.

#### **What will happen to the results of the study?**

The results of this study will be used for academic purpose only and will be published in the final dissertation and may be used in other related academic publications.

#### **What do I get for participating?**

Participation is voluntary and will involve no costs or payments to you.

#### **What if I am not happy with how the research was conducted?**

Please contact Pat Dixon dixonap@aol.com

### Appendix 3: Informed consent form

#### Research data management: a case at the Loughborough University Library

##### INFORMED CONSENT FORM

The purpose and details of this study have been explained to me. I understand that this study is designed to further scientific knowledge and that all procedures have been approved.

I have read and understood the information sheet and this consent form.

I have had an opportunity to ask questions about my participation.

I understand that I am under no obligation to take part in the study.

I understand that I have the right to withdraw from this study at any stage for any reason, and that I will not be required to explain my reasons for withdrawing.

I understand that all the information I provide will be treated in strict confidence and will be kept anonymous and confidential to the researcher unless otherwise stated.

I would like to keep my identity anonymous YES NO (please underline)

I agree to participate in this study.

Your name	
Your signature	
Date	

## Appendix 4: Interview protocol

### **Introduction:**

Thank you for dedicating me your time and taking part in this study.

### **Recording and confidentiality**

With your permission I would like to record this interview (audio). What you say in this interview will be transcribed and sent to you for authorisation.

**Purpose:** The purpose of this study is to explore how the Loughborough University Library is embracing research data support services.

More specifically I would like to look at **requirements, roles and responsibilities** of some stakeholders in the implementation of research data management.

**Overview:** Firstly I would like to ask you about how research and data practices are currently supported. Also I would like to look at services that will be implemented, and finally talk about your needs and requirements.

Do you have any other comments about what we have discussed, or about the research as a whole?

### **Closure:**

Thank you for allowing me to talk to you today, it has been very interesting to listen to your views. I will email the interview material, which will be used only after your permission and your amendments. Thanks again, I am very grateful for your help.

## Appendix 5: Email interview

1. Can you explain me briefly what is your role at Loughborough University
2. Could you describe your role and responsibilities in supporting research community and research data management (RDM)?
3. Could you describe the services that the library provides to researchers focusing on those that are related to research data management?
4. What are the gaps and weaknesses in the provision of current research services?
5. In your opinion what other services would help manage research data more effectively?
6. Are there any services planned to be set up?
7. In your opinion are there other stakeholders who should be involved in the implementation of research data management?
8. Are there any partnership which should be establish to enhance the provision of research data support?
9. Are there any services, training or tools that are required to support you in providing research support services?
10. In your opinion is the library workforce equipped to support research data management?
11. Who is or could be best qualified in supporting research data management?
12. In your opinion what is the role of providing support services to researchers and supporting research data? In your opinion why research data matters to you and to the library?
13. Can you briefly introduce some of the main challenges when supporting researchers in research data management? Can you identify three key issues to be addressed what would they be?
14. What are the key elements to build the data management support at Loughborough University?

## Appendix 6: Interview guide

Questions	Prompts
1. Could you describe your role and responsibilities and those of your department in supporting research data management? Can you begin by describing how you place yourself with relation to RDM?	RDM Working group Services
2. What are the gaps and weaknesses in the provision of current research services?	Researchers lack of recognition of the library's role No coherent, central approach Lever from ESRC UK Data Archive
3. In your opinion what other services would help manage research data more effectively?	RDM roadmap Training, education New hires
Are there any services planned to be set up?	IT service
4. In your opinion are there other stakeholders who should be involved in the implementation of research data management?	'Champions' researchers, Library, Computing services, IT Office, Research Office, Enterprise Office, Intellectual Property Office, Information Science department More researchers Institutional priorities
5. Are there any partnership which should be establish to enhance the provision of research data support?	facilities management external partners
6. Are there any services, training or tools that are required to support you in providing research support services? Is there anything that could support you?	International standards, expectations resources, time Institutional message DCC roadshow with best practice and exemplars from other universities
7. In your opinion is the ... workforce equipped to support research data management?	Awareness
8. Who is or could be best qualified in supporting research data management?	RO, Library Combination One stop shop
9. In your opinion what is the role of providing support services to researchers and supporting research data? Why research data matters to you and to the ...?	RC, Good for research institution Awareness of researchers
10. Can you briefly introduce some of the main challenges when supporting researchers in research data management? Can you identify three key issues to be addressed what would they be?	Lack of assessment of the value of research data Ensure that RD are reusable for future researchers Storage, education, definitions Cost? Lack of recognition
11. What are the key elements to build the data management support at Loughborough University?	Skilled staff, education Infrastructure Recognition of researchers senior management